"GEOGRAPHY AND THE PROPERTIES OF SURFACES" SERIES

Paper Number Twenty One

GEOGRAPHY AND AN EXISTENCE THEOREM

A Cartographic Computer Solution to the Localization on
a Sphere of Sets of Equal-Valued Antipodal Points for
Two Continuous Distributions with Practical Applications
to the Real Earth

by Stephen E. Selkowitz
Harvard College '70

HARVARD PAPERS IN THEORETICAL GEOGRAPHY

Laboratory for Computer Graphics and Spatial Analysis

Harvard Center for Environmental Design Studies
The Graduate School of Design
Harvard University, Cambridge, Mass.

5 August 1968

PREFACE

The earth's surface may, to a first approximation,
be regarded as spherical in geometrical terms and, in
topological terms, it may be considered as a single
continuous closed surface.

In physical geography the sphericity of the earth
has been recognized and assigned a role. Spatial pre-
diction on this basis has been so very successful as to
have been commonplace and no longer regarded to be
spatial prediction as such. Thus, the statements about
sunrise and sunset, length of day, seasonality, celestial
appearances and navigation, and the like which depend
in part upon the recognition of sphericity of the earth's
surface are regarded as certainties and the predictive-
ness involved is lost sight of.

The topological nature of the surface has not re-
ceived as much attention and the requirements that this
places upon the theories of spatial process in the physical
geography of global patterns of circulation have not been
recognized, to the disadvantage of those theories.

Imagine now two spatially continuous distributions
covering the surface of the earth. In physical geography
we might consider temperature and pressure distributions.
In socio-economic terms the population potential field
and the income potential field serve as examples. Given

only that there are two continuous distributions, A and
B, over the earth's surface, there will always be, as
Steinhaus has noted, at least one pair of antipodal points
having both the same value for A and the same value for
B. Through time the spatial distributions involved may
change and the given antipodal points initially so
related as above may be no longer. However, another
such pair will then exist and there will have been a
shift in position. This is the consequence of the theorem
proposed by Ulam and proved by Borsuk that if a sphere
is folded and distorted (but not torn) so as to be made
to lie flat, there is necessarily one pair of antipodal
points from the sphere which come to lie upon each other
in the new situation.

The proof of this existence theorem, however, pro-
vides no means for finding specific locations for the
equal-valued antipodal points for the two continuous
distributions. This concerns us.

Theoretical geography is a science of earth loca-
tion and spatial relations. It describes, classifies,
and predicts locations in the spatial sense. Cartographics
stands to geographical science as graphics does to science
generally. Mapping is a general mathematical concept
in the theory of sets. Cartography is the geographical
example in the application of this concept. That which
is ordinarily called a map by geographers and laymen is

11

technically "a graphical image of a mapping." Whatever
useful roles graphics plays in science generally also
can be claimed for cartography with relation to geograph-
ical science. This is especially true now that geographers
increasingly employ geometry as an appropriate vehicle
to carry their discipline.

There is already, of course, an accumulated stock
of knowledge and experience concerning maps as stores of
spatially ordered information. We hope, however, to
examine the expanded roles that mapping (especially
computer-assisted mapping) seems well-suited to play in
the sciences viewed from the standpoint of theoretical
geography and in the disciplines employing its models
for decision-making purposes.

We recognize yet another role for maps. In the
solution of certain problems for which the mathematics,
however elegantly stated, is intractable, graphical
solutions (approximations) are possible. This is
especially true with regard to "existence" theorems.
There are many cases in which the graphical solution to
a spatial problem turns out to be a map in the full
geographical sense of the term, map. Thus, a map is
the solution to the problem. Computer-achieved maps
now are possible in this connection whereas previously
computational burdens rendered them virtually unattain-
able.

Mr. Stephen Selkowitz, Harvard College '70, under-
took to find a method for locating specifically desired
sets of antipodal points on a sphere using both hypothet-
ical data and information concerning patterns on the real
earth and within the constraints of the SYMAP (computer
mapping) program.

> William Warntz
> Professor of Theoretical Geography
> and Regional Planning
>
> Harvard University
> Cambridge, Massachusetts

20 July 1963
Cape May, New Jersey

GEOGRAPHY AND AN EXISTENCE THEOREM -- A Cartographic Computer Solution to the Localization on a Sphere of Sets of Equal-Valued Antipodal Points for Two Continuous Distributions with Practical Applications to the Real Earth

One can discuss the existence of sets of antipodal points on spheres without concern for their specific spatial locations. A mathematician working with a certain existence theorem, for instance, may be primarily interested in whether or not a given phenomenon concerning the pairing of antipodal points exists; the frequency of occurence and location may be beyond the scope of the proof. Consider the following example: given any two continuous functions, f and g, mapped over the surface of a sphere, there must exist at least one set of antipodal points at which function f is equal and function g is equal. The proof of the theorem gives no information about how many other such pairs of points exist or where, on the surface of the sphere, any of such sets is located. One might expect that geographers and others would be interested in questions of this nature, however. This paper is an attempt to answer these questions by developing a procedure to find spatial solutions to the previously stated existence theorem with the aid of and existing SYMAP computer mapping program. (See paper No. 3 of the Harvard Papers in Theoretical Geography, "Implicit Map Projections in Computer Print-Outs," and paper No. 15, "A Two-Dimensional Interpolation Function for Computer Mapping of Irregularly Spaced Data," for a

description of this program.) Stated simply, the problem
is to find on a sphere all pairs of antipodal points on
the sphere at which the values for some continuous function
f are equal <u>and</u> the values for some continuous function g
are equal. Of course the values for each of these functions
may differ from set to set and for real phenomena the
actual values depend upon the units chosen.

The first step in rendering the problem suitable to
a computer solution was to change the problem domain
from the surface of a sphere to a plane. There are any
number of methods of projecting a sphere onto a plane
but the possible choices are limited depending upon
several requirements of the projection. Most importantly,
the projection must facilitate the comparison of function
values at antipodal points on the sphere. Also, since
the continuity of the functions is crucial, the projection
should make a minimum of cuts in the surface of the
sphere and those should be of a regular nature. A set of
two hemispherical projections was chosen as best suited
for these requirements. If the sphere is split into two
hemispheres and each is projected onto a plane lying
parallel to the cut, a rotation of one of the hemispher-
ical projections by 180° will bring all sets of antipodal
points into alignment. If, however, as is usually the
case, the projections are made from antipodal points to
tangent planes, one of these must be inverted before
rotation to bring all antipodal points into alignment.

Now consider what the locus of equal-valued antipodal

A, B
ARE ANTIPODAL POINTS

HEMISPHERICAL
MAPS

1. 2.

3. 4.

•A

B•

ROTATE
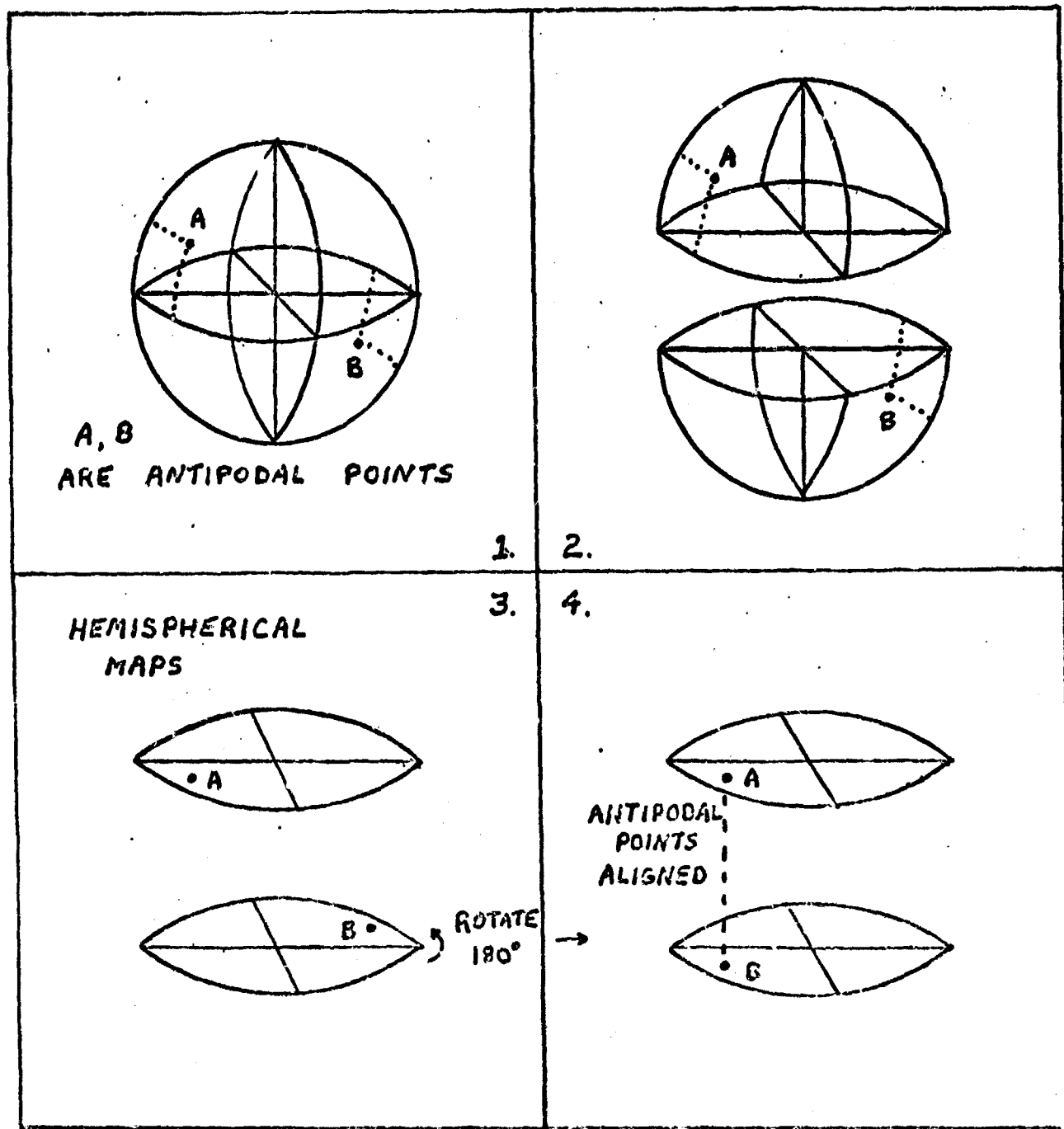180°

ANTIPODAL
POINTS
ALIGNED

•A

•B

Figure 1. Simple Transformation to Align Antipodal Points

points might look like on the hemispherical projections.
The problem is further simplified by assuming the globe
to be the sum of many great circles and then examining
the behavior of the function along any great circle.
The conclusion is that there is at least one set of
antipodal points where the function has equal values on
every great circle. The proof is simple. Since the
function is continuous on the whole sphere, it must be
continuous on the great circle. Choose a diameter of
the great circle: its endpoints are antipodal points.
Rotate the diameter along itself through 180°, subtracting
the value of the function at one end of the diameter from
the value at the oposite end (the antipodal point). If
the initial result of the subtraction be some number A,
when the diameter has completed its arc the result of
subtraction would be -A. Since the value went from A
to -A and, because the result of subtraction of contin-
uous functions is continuous, at some set of points the
result is zero. At that set of antipodal points, the
function has the same value. Thus any great circle has
at least one set of equal antipodal points.

Superimpose the two hemispherical projections of
the globe so that each point on one hemisphere is aligned
with its antipodal point on the other hemisphere. The
result is a circular map, each of whose points represents
a pair of antipodal points. Every diameter of this circu-
lar map is a great circle of the original sphere. (There
are, of course, other great circles besides these diameters.

The characteristic of all great circles on these circular
maps is that they intersect the circumference of the map
at diametrically opposite points.) It was previously
shown that every great circle (each diameter on this map)
will have at least one set of equal antipodal points. If
a whole series of diameters, each with its set of equal
antipodal points, be drawn on the map, a reasonable assump-
tion, by virtue of the continuity of the function, is
that the sum of these points will form a curve of some
sort. That is, in the limit of an infinite number of
diameters, one would expect the locus of equal sets of
antipodal points to be a curve. Note that where the curve
intersects the circumference (it must, since the circum-
ference is a great circle) it must also reappear at the
diametrically opposite point. This is because each point
on the circumference not only represents a set of antipodal
points, but is the antipodal point of its diametrically
opposite point on the circumference (i.e., the circum-
ference duplicates a set of antipodal points).

Thus the locus of equal-valued antipodal point sets
should be a closed curve (only the particular projection
used makes the locus appear to have endpoints). There
can also be a single isolated set of equal-valued antipod-
al points. These events, however, represent somewhat
special cases and will be discussed later.

One might also ask what is the minimum number of
such closed curves and what is their location? This
minimum locus demands that every conceivable great circle
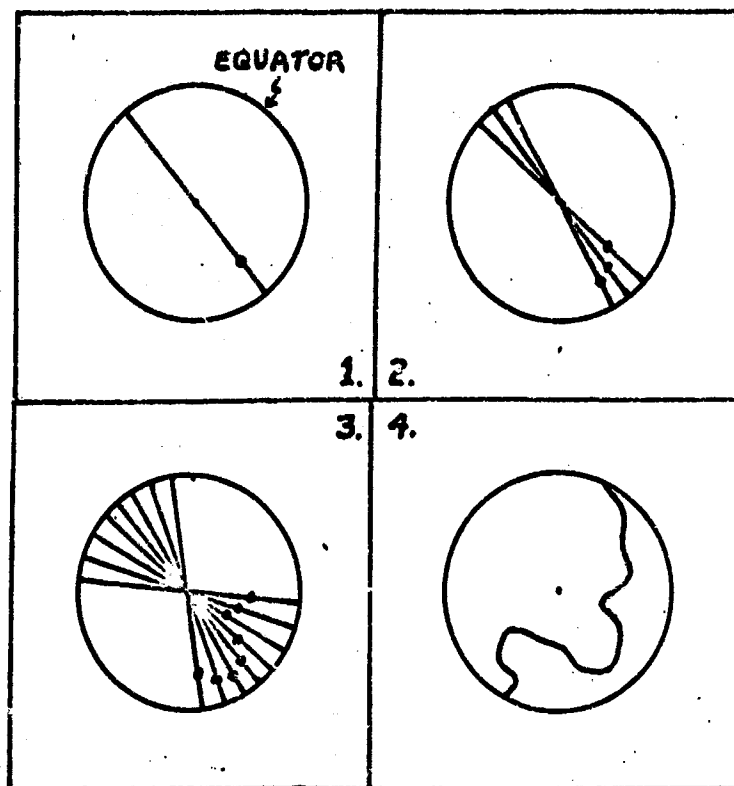intersect it at least once. Intuitively, any line

Figure 2. The locus of sets of antipodal
points with equal values is
generated point by point.

connecting two opposite points on the circumference of
the circular map will intersect, at least once, every
great circle which can be drawn. This follows from the
previously stated observation that all great circles
begin and end at pairs of diametrically opposite points
on the circumference of a circular map. Thus the sim-
plest locus of equal-valued antipodal points might be
a diameter or half of the circumference of the circular
map, or more generally, the projection onto the circular
map of any great circle on the sphere. This is a
plausible conclusion since every great circle on a sphere
intersects every other great circle at one set of antip-
odal points.

If the minimum locus of equal-valued antipodal
points for one function is any line connecting two
opposite points on the circumference of the circular map,
the same must be true for any other function. The impos-
sibility of connecting each of two sets of diametrically
opposite points on the circumference with lines which
do not intersect at least once is an intuitive affirma-
tion of the original existence theorem. With these
expectations as to the form of the solution, the solution
procedure can now be developed.

A brute force solution is described first. For
each of two functions, use the SYMAP program to produce
contour maps (on hemispherical projections as base maps)
of both hemispheres. Taking one function at a time,
align the two hemispherical contour maps so that antipodal

points are aligned and mark the locations where the same contour levels cross. This will produce not the line as hypothesized before, but a string of polygons whose size will depend on the size of the contour levels. (This will be discussed in more detail later.) After carrying out the same procedure for the second function, it is a simple matter to compare the two loci of equal-valued antipodal points to see where they intersect.

The first step is to produce contour maps using the SYMAP program. But even before a base map can be plotted, a particular hemispherical projection must be chosen. Since the accuracy of the SYMAP interpolation is dependent upon the area over which the interpolation occurs, a hemispherical projection which preserves the area of the sphere (although this will mean some distortion of linear distances) is an appropriate choice.

While the theorem concerns functions mapped on any sphere, the examples can be restricted to a globe of the earth without loss of generality. This has the advantage of providing convenient terminology, such as north and south, and a grid system of latitude and longitude for locations as well. Thus, for conven' ice onl , center the projections on the North and South Poles; the circumference of the previous circular maps is, then, the equator.

This map projection is called a polar azimuthal equal area projection. The radius of a latitude line on the map is set equal to the length of the chord from

the pole to the same latitude line on the globe. Thus
the radius of the projected map is $\sqrt{2}$ times the radius
of the globe but the area of the map is equal to the
area of one hemisphere of the globe. The greatest
linear distortion will occur at the equator where the
circumference of the map is $\sqrt{2}$ times the circumference
of the equator on the globe.

A globe of 4 inch radius was chosen from which to
project the polar equal area maps (the radius of the
map was thus 5.66 inches). These numbers were chosen
so that the map would fit on one panel of computer
printout. Choosing the number of data point locations and
spacing was a less rational process. The data points
had to be evenly spaced and close enough together to
avoid missing phenomena of the functions. A symmetrical
arrangement was an aid to plotting data point locations
and facilitated rotating the maps so that the antipodal
points were aligned. The resultant base map for the first
series of maps contained 64 data points within the equator
spaced as shown on Figure 4.

Functions generated with random numbers were used
for the first attempts at perfecting a method to solve
the problem to avoid any singularities or pecularities
which some naturally occuring function might have. By
assigning random numbers from 1 to 99 to every data
point and interpolating with SYMAP, the result was a
contour map of a continuous function, with one important
exception, i.e., at the equator. If random numbers were

0°

15°

30°

45°

60°

75°

90°

R

HALF OF POLAR
AZIMUTHAL EQUAL
AREA PROJECTION
FROM GLOBE

$R = \sqrt{2}\, r$

90° 75° 60° 45° 30° 15° 0°

CROSS SECTION
OF NORTHERN
HEMISPHERE
OF GLOBE
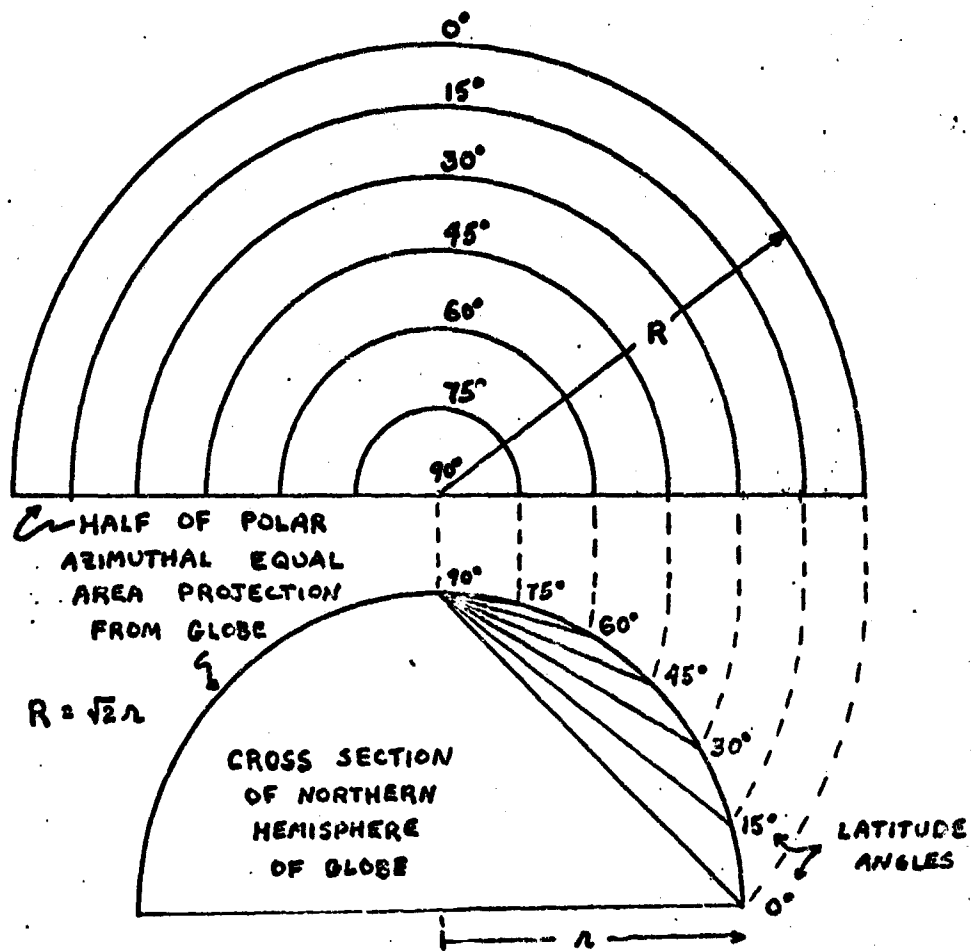
LATITUDE
ANGLES

r

Figure 3.   Projection of Polar Azimuthal Equal Area
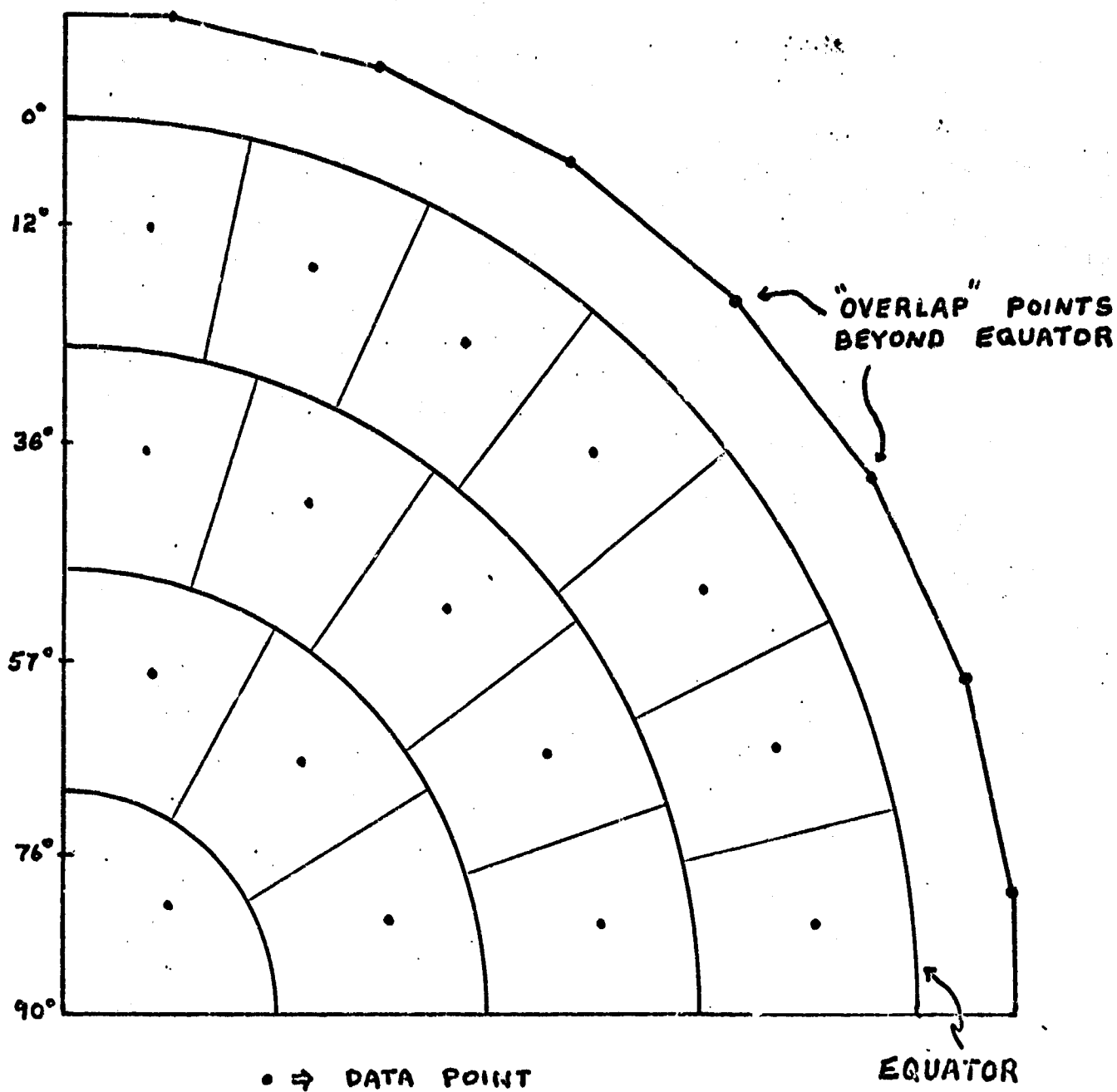
Map from the Globe

Figure 4.   Data Point Locations on one Quadrant of Base Map

assigned to data points in each hemisphere and interpola-
ted out to the equator, the function values at the
equator of both hemispheres were generally not the same.
But, the functions must be continuous over the whole
sphere for the theorem to hold. To provide continuity
at the equator, 28 data points were taken from one
hemisphere and added to the other hemispherical map in
the region beyond the equator. For example, the ring
of data points closest to the equator in the southern
hemisphere was placed beyond the equator in the
appropriate position in the northern hemispherical map.
The computer would then interpolate across the equator
on both maps and provide the same values along the equa-
tor in both hemispheres. Subsequent maps confirmed
that values along the equator agreed to within 1 or 2
character locations from one hemisphere to the other.
With these 28 points in the "overlap" region, there was
a total of 92 data points per map.

After a set of trial maps with five contour levels
confirmed the validity of the solution to the continuity
problem at the equator, a set of two hemispherical maps
for each of two different randomly generated functions
was produced with a radius of 11.32 inches (two panels
of computer printout). The maps were done in 10 levels
with white contour lines between the levels. These had
been omitted from smaller maps because they constituted
a large percentage of the available character locations.

The procedure for finding the locus of equal-valued
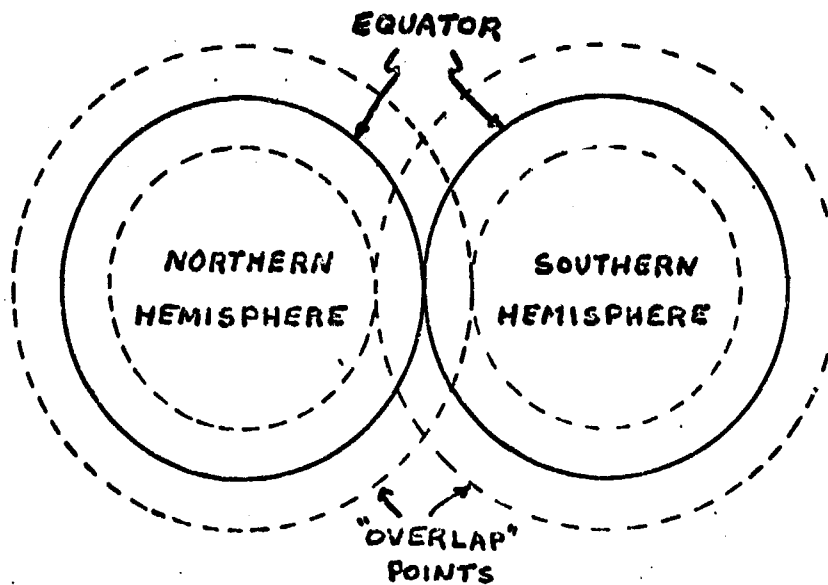antipodal points is as follows. For each function,
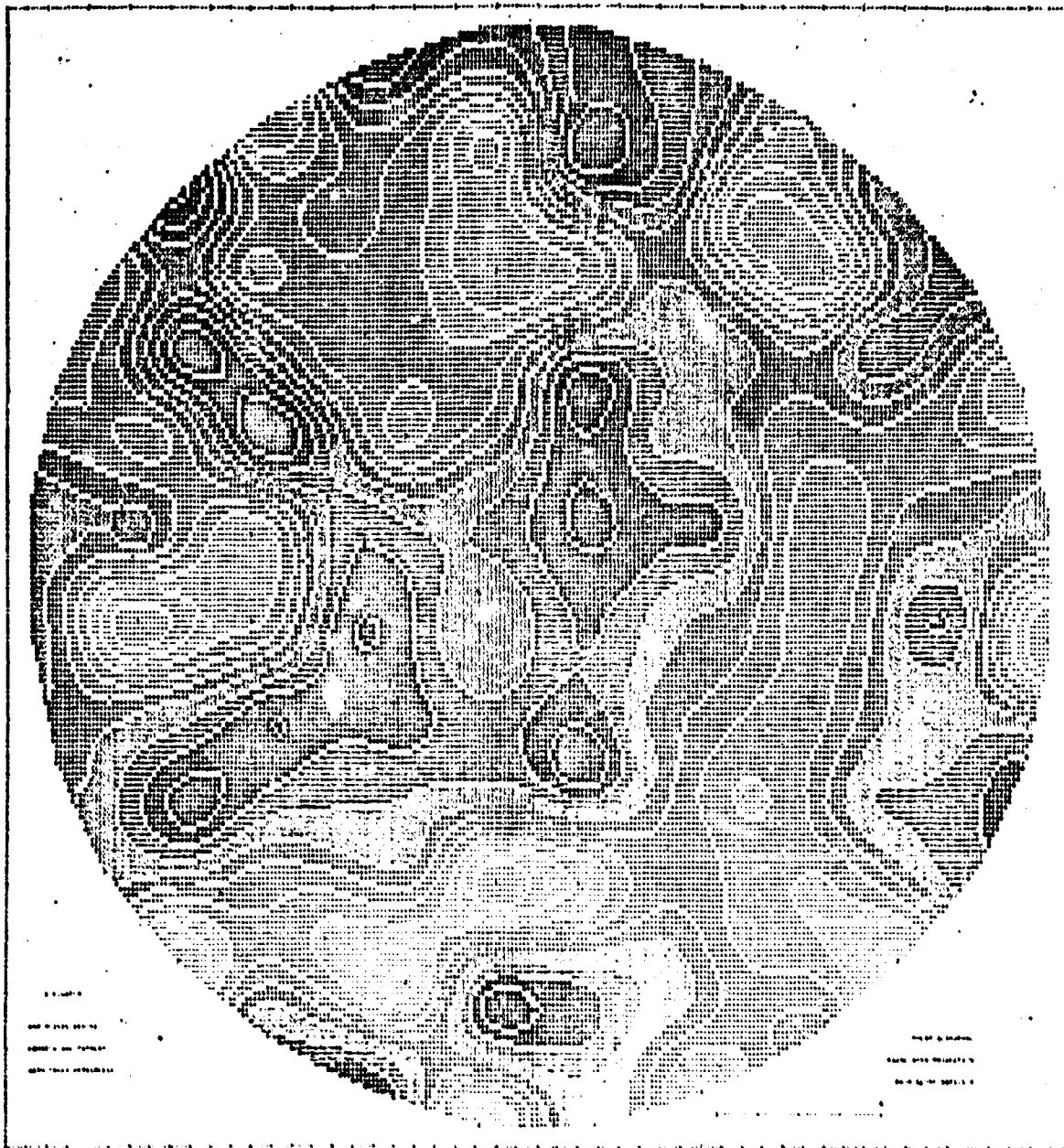
Figure 5.   Generation of "Overlap" Points

Figure 6. Random Function Mapped on one Hemisphere

place a sheet of clear plastic over one of the hemis-
pherical maps and trace all contour levels with a grease
crayon, labeling the contour levels with appropriate
values (1-10). Then place this tracing over the other
hemispherical map so that antipodal points are aligned
and mark the regions where contour levels of the same
value cross. The result for one of the randomly genera-
ted functions is illustrated below.

As noted earlier, the locus of equal-valued antip-
odal points shows not as a series of closed curves, but
rather as a string of polygons. The reason for this is
that a continuous funciton has been approximated with a
step function and all points in the same contour level
are taken as equal in value when in fact they represent
a range of values. As the number of contour levels in-
creases, the range of each contour level will shrink.
Because the functions are continuous, in the limit of
an infinite number of step functions, each of vanishingly
small range, one would expect the string of polygons to
shrink to a series of closed curves. This process is
illustrated below where the number of contour levels has
been doubled. The area of the polygons shrinks approx-
imately by a factor of two. Note that only when the lines
dividing the same contour levels intersect (point A, for
example, on Figure 8a) are there equal-valued antipodal
points. Points within the polygons are sets of equal-
valued antipodal points only to within the value range of
the contour level.

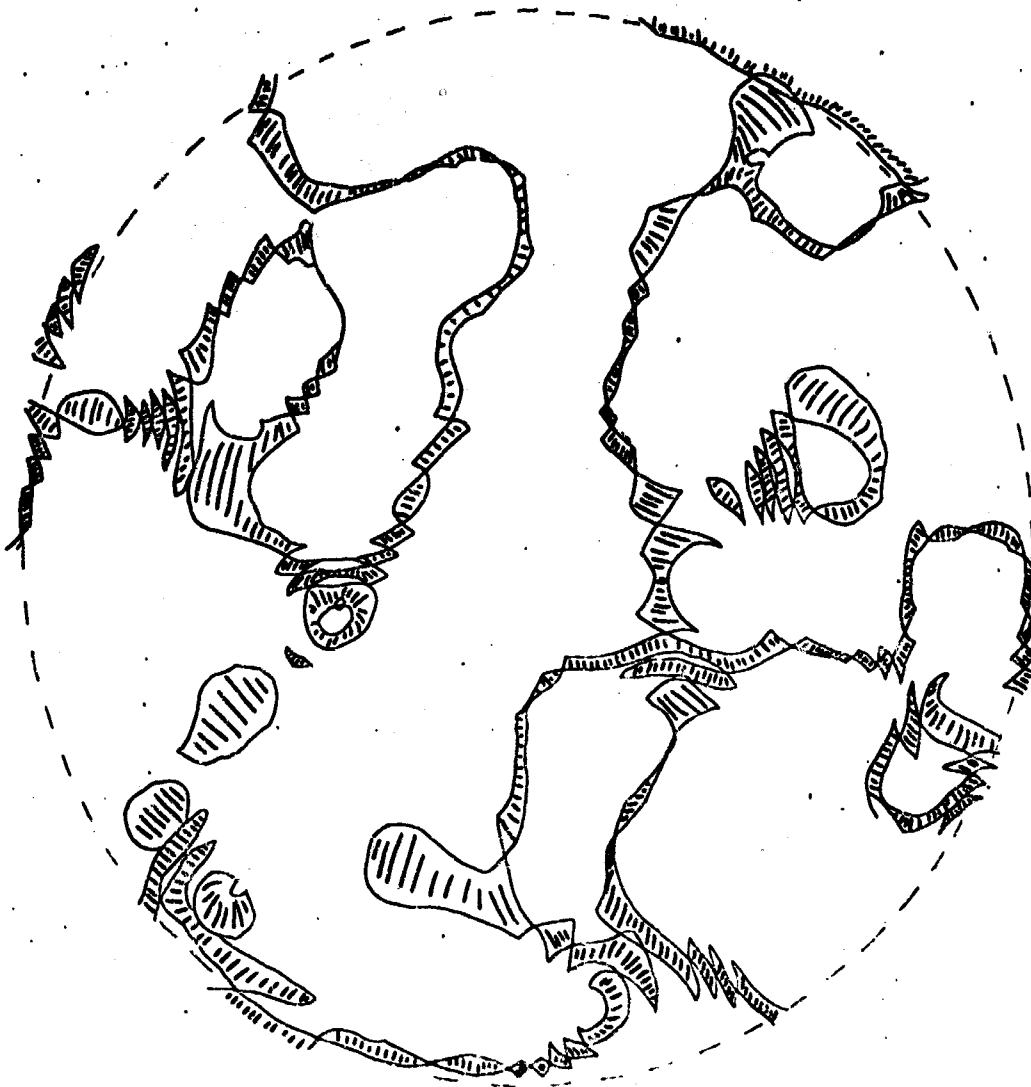The strings of polygons can be approximated with a

Figure 7. The shaded area is the locus of sets of antipodal points with equal function values for one random function.
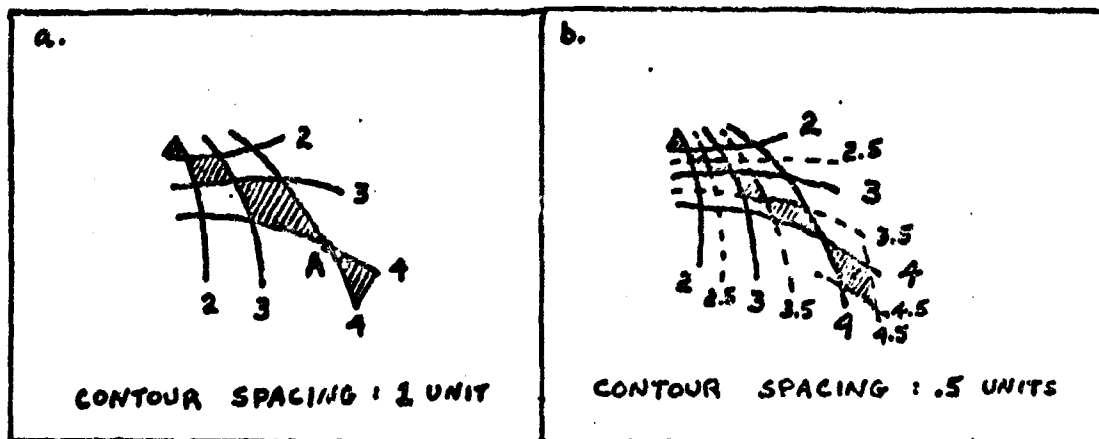
Figure 8.   Total Polygonal Area as Function of
Contour Intervals

series of lines. Although the precise position of the line of sets of equal-valued antipodal points within the polygons is often uncertain, there is a series of points where contour lines intersect through which the locus must pass. But, there is still considerable uncertainty about where the locus of equal-valued points is located when the polygons are large in size and irregular in shape. (Figure 9b)

Thus it is with uncertainty that the strings of polygons are reduced to a series of lines in Figure 10. The loci of sets of equal-valued antipodal points for both functions are drawn on the same map. Within the limits of the aforementioned uncertainty, each time one of the loci vanishes at the equator it reappears at the opposite point. There are about 24 points of intersection where both randomly generated functions have equal values at sets of antipodal points.

This method gives a satisfactory, though not ideal, approximation to the problem. By using the computer maps merely as stores of spatially ordered data, problem solving capabilities have not been utilized to the fullest extent. The actual manipulation of the data was a human effort and a tedious one at that. As a refinement of this preliminary technique, the computer can be involved in more than the initial display of the data -- it can manipulate the data in some way relevant to the solution of the problem and then graphically display the results of the manipulation.
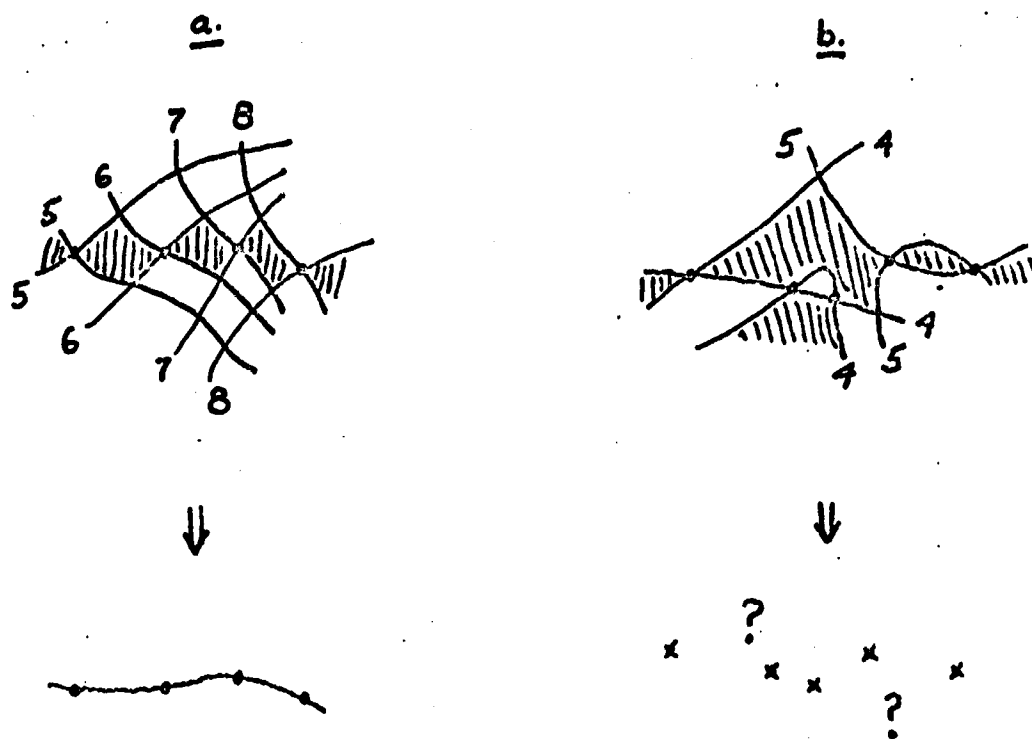
Figure 9. The process by which strings of polygons are reduced to lines is subject to great uncertainty.
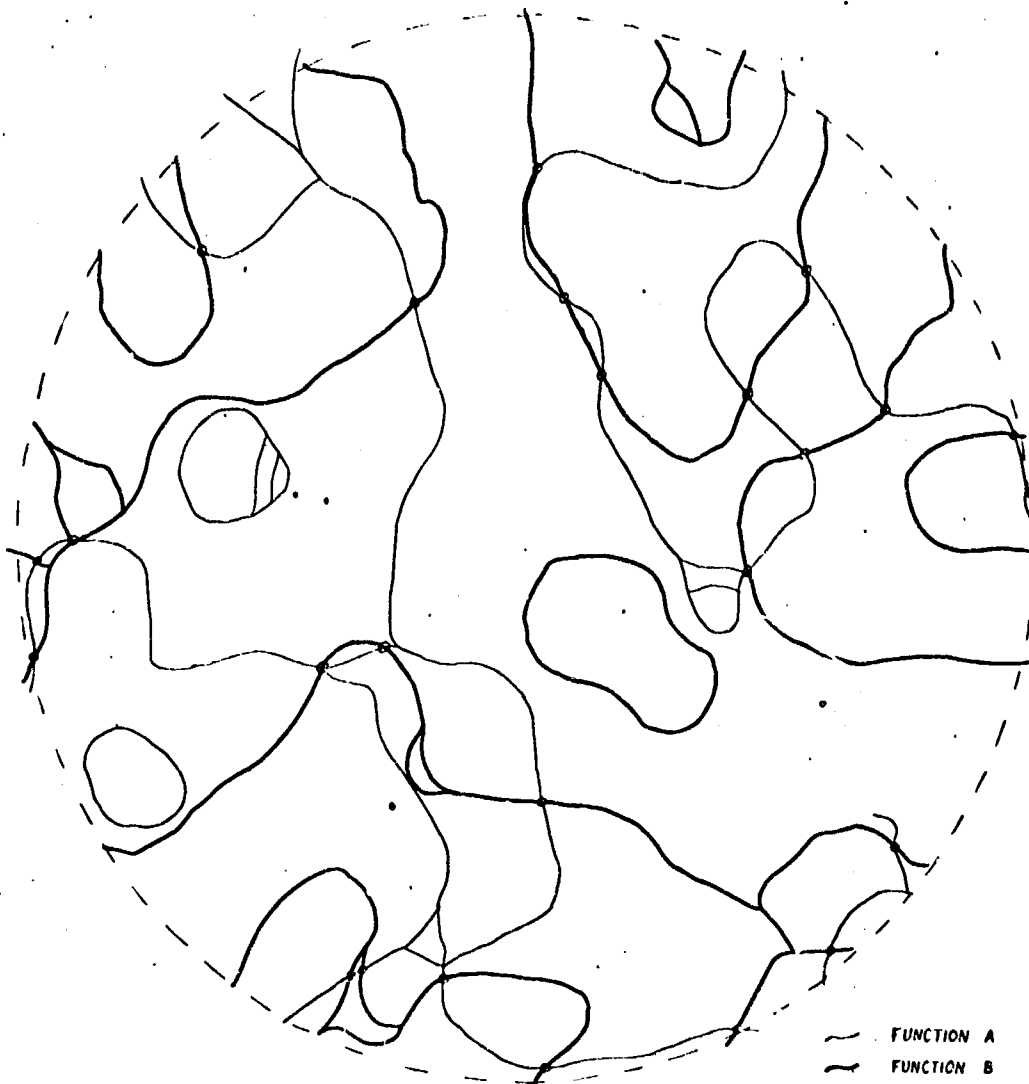
Figure 10. Loci of Sets of Equal Valued Antipodal Points for
Both Random Functions

note: The intersections are circled.

The first step is to eliminate the tedious work involved in tracing contour levels and searching for overlapping contours, a process that accumulates inaccuracies with multiple tracings, etc., and is susceptible to errors of omission on large complicated maps. The process of comparing the function values at sets of antipodal points is analogous to subtracting the function value at each point from the value at its antipodal point. The result of this process over the whole hemispherical map is called the "difference function", a continuous function. Every point on the difference function mapping represents two points on a globe, a point on one hemisphere and its antipodal point on the other hemisphere. When the difference function is zero, the function, at that set of antipodal points, has equal values. If, at some point, the difference function has a positive value and, at a nearby point, the difference is negative, the value of the difference functic. must pass through zero on any path joining the two points, a consequence of the continuity of the function. The set of points where the difference function takes on the value of zero is the locus of equal-valued sets of antipodal points.

With the use of the computer, the thousands of subtractions necessary to calculate the difference function at every character location is not a difficult feat. A subroutine can be programmed into the SYMAP program which allows the computer to store on tape the function values at every character location for two different functions

on the same base map (actually the same function on two different hemispheres) and then subtract function values at each character location. The computer could then be instructed to print symbolism only at those locations where the result of subtraction was zero, or since a difference of exactly zero would be unlikely, at all locations where the result of subtraction was between $\pm\delta$, where δ is a small number which would vary with each function.

As first approximation to subtraction at every character location, several difference functions were generated by computing the value of the difference function only at data points (by subtracting the value at a data point in the northern hemisphere from the value at the antipodal point in the southern hemisphere) and then interpolating over the whole map to produce values for the difference function at every character location. This was considerably simpler and less costly than writing a subroutine, debugging it and then subtracting at about 8000 character locations.

Difference functions were mapped in 10 levels with no white spacing for contour lines for each of the two randomly generated functions. Had there been many data points or large numbers to subtract, the values of the function at corresponding antipodal points in each hemisphere could have been entered in a data bank and the subtraction performed by the computer. Since there were only 92 subtractions of two digit numbers, it was
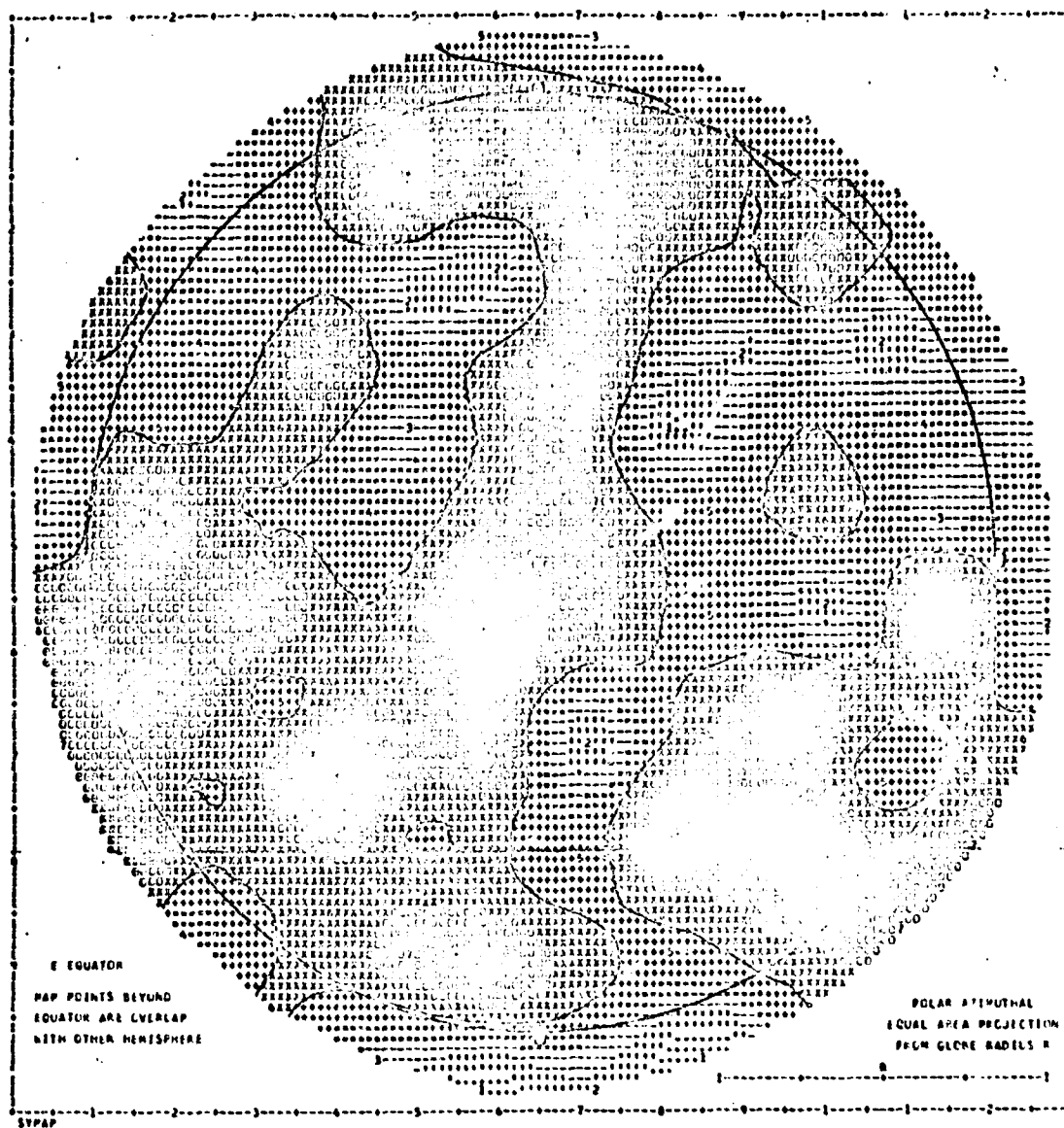
Figure 11.   The zero contour for the difference function is
             marked between levels 5 and 6.   The equator is
             also indicated.

just as simple to do the subtractions by hand and use
the results as data point values of the difference
function mapping.

The printout of the 0 level for one of the difference
functions is shown in Figure 11. For this function, the
0 level corresponded to a contour line between levels 5
and 6 and thus could be easily plotted. For the second
function, the 0 level contour fell in the middle of a
contour level so there was a certain amount of guesswork
as to where it actually was. The resultant locus of
equal-valued sets of antipodal points for each of the
two functions was in general agreement with the previously
plotted results.

Having now established two preliminary methods for
finding the location of the points that the existence
theorem says must exist, I decided to do further work
with real functions as opposed to the functions generated
with random numbers. Two completely unrelated functions
could have been used and the theorem would, of course have
held, but if for real phenomena related functions were used
there could have been some significance to the spatial loca-
tions of the sets of antipodal points at which each function
had equal values. The choice of functions was governed
mainly by what was available. I desired two related
functions mapped over the whole earth with a minimum of
10 contour levels and preferably displayed on a polar
equal area projection. Two maps in the series "World
Maps of Climatology" by H.E. Landsberg, H. Lippman, K.H.

Paffem, and C. Troll, 1963, met the first two requirements
but not the last. The two functions were total sunshine
received by the earth in hours per year (hereafter called
the Sunshine function) and average annual energy from
sky radiation in kcal/sq. cm/year (the Energy function).
Were it not for atmospheric factors, etc., both functions
would have depended on latitude only and the functions would
have had equal values at every set of antipodal points.
Because of atmospheric and other disturbances, the maps
showed great deviations from a latitude dependent phenomena.
It might be revealing to note the sets of antipodal
points at which the factors which made one function have
equal values also made the second equal-valued. This was
just a restatement of the original theorem.

Since the Energy and Sunshine functions were not
available on hemispherical maps, the contours were redrawn
on a polar projection to facilitate finding the function
values at data points. This was not really necessary --
the location of each data point on the polar projection
could have been determined in terms of latitude and
longitude and then the function value at that point on
the original map could have been found. One could also have
derived equations to convert from latitude and longitude
on a sphere to SYMAP coordinated on the base map. The
equations are:

$$\text{Row} = y_0 - \frac{8(2+\sqrt{2})R\cos\phi}{1+\sqrt{2}+\sin\phi}\cos\theta, \quad \text{Column} = x_0 - \frac{10(2+\sqrt{2})R\cos\phi}{1+\sqrt{2}+\sin\phi}\sin\theta$$

for a polar azimuthal equal area projection from a globe
of radius R where ϕ = latitude, θ = longitude and (x_o, y_o)
are coordinates of the center. (θ is measured counter-
clockwise from the top of the map.) But this would not be
really useful unless one were willing continually to change
data point locations.

The loci of equal-valued sets of antipodal points
for the Energy and Sunshine functions were determined
first by the laborious hand tracing technique and then by
using the difference functions. As before, while there
was some uncertainty in the exact position of the equal-
value loci with the hand tracing method, every time the
locus intersected the equator it reappeared on the opposite
side as it should have. The difference function approach
lacked this virtue although when the zero contour fell
between two contour levels there was no guesswork as to
the location of equal-valued sets of antipodal points.
It appeared that the difference function missed some of
the 0 contours because the data point spacing was too wide.
As a refinement of this technique, reconsideration of
the possibility of subtracting at every character location
was in order. An examination of the interpolation process elim-
inated this process for the following reasons.

Subtraction at every character location will produce
a significant improvement in the accuracy only if the
function values to be subtracted have been computed to
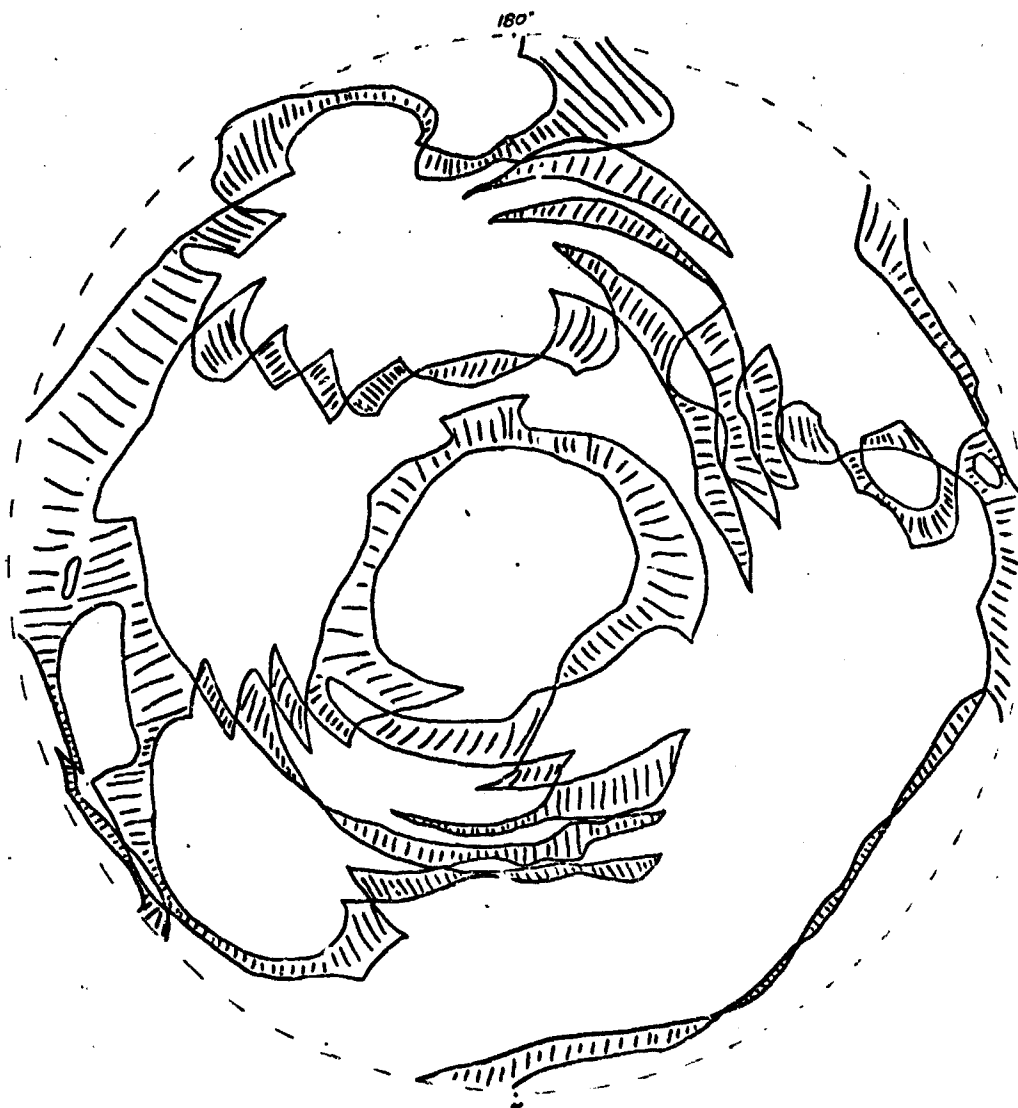great accuracy at every character location. The SYMAP program

Figure 12a. The shaded area is the locus of sets of antipodal points with equal function values. The function is the total annual energy incident upon the earth's surface.
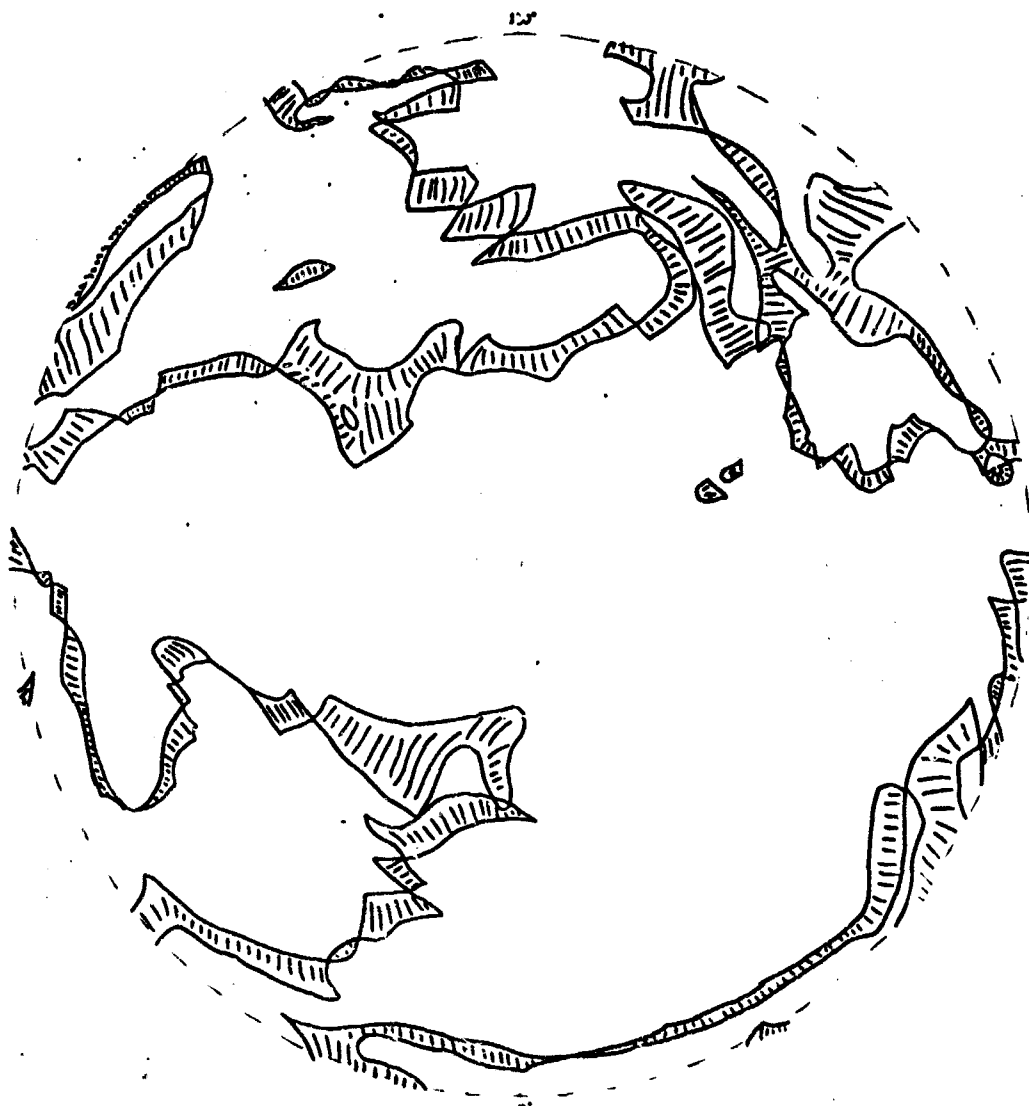
Figure 12b. The shaded area is locus of sets of antipodal points with equal function values. The function is the annual sunshine duration at the earth's surface.
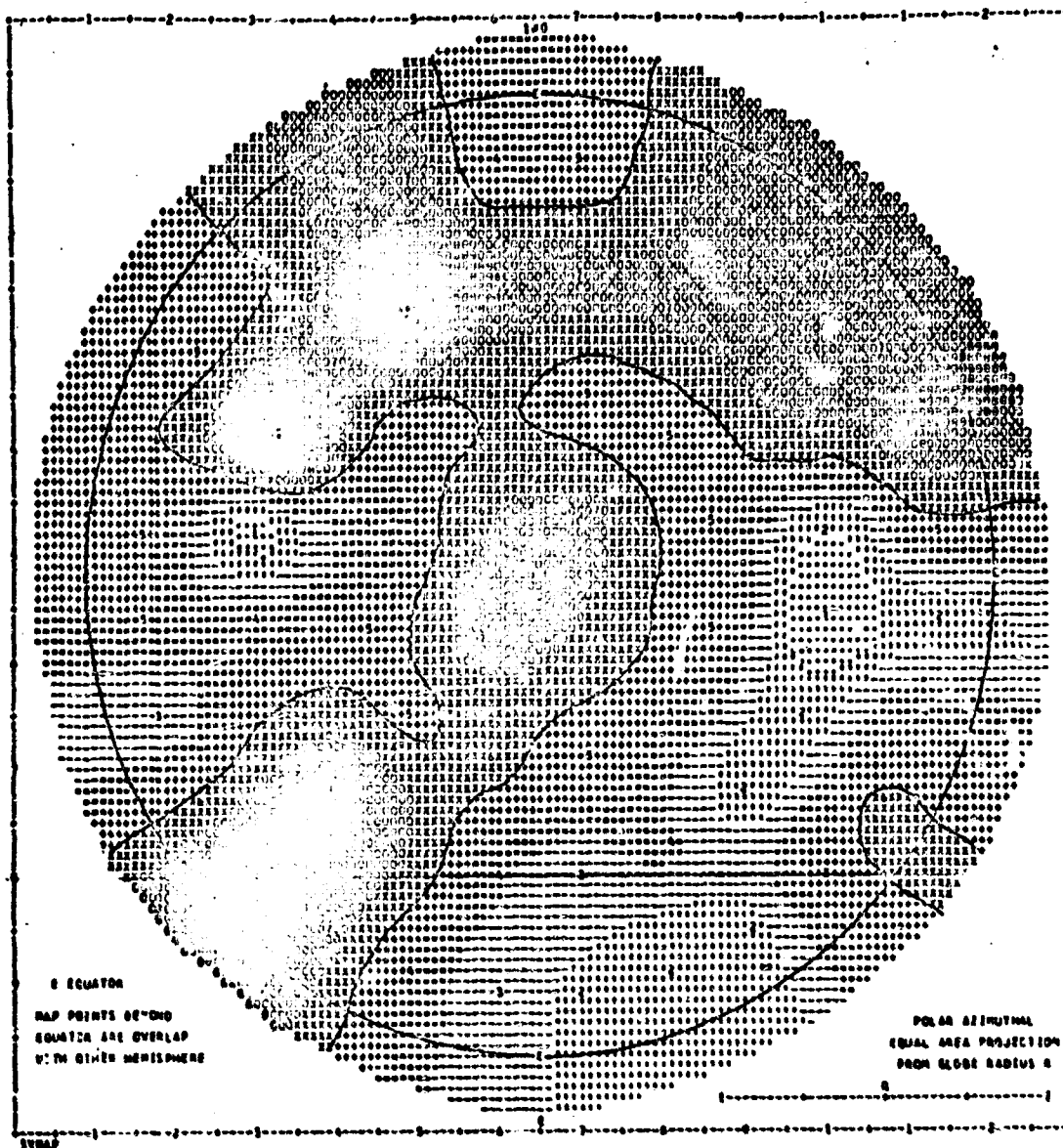
Figure 12c. The zero contour for the energy difference
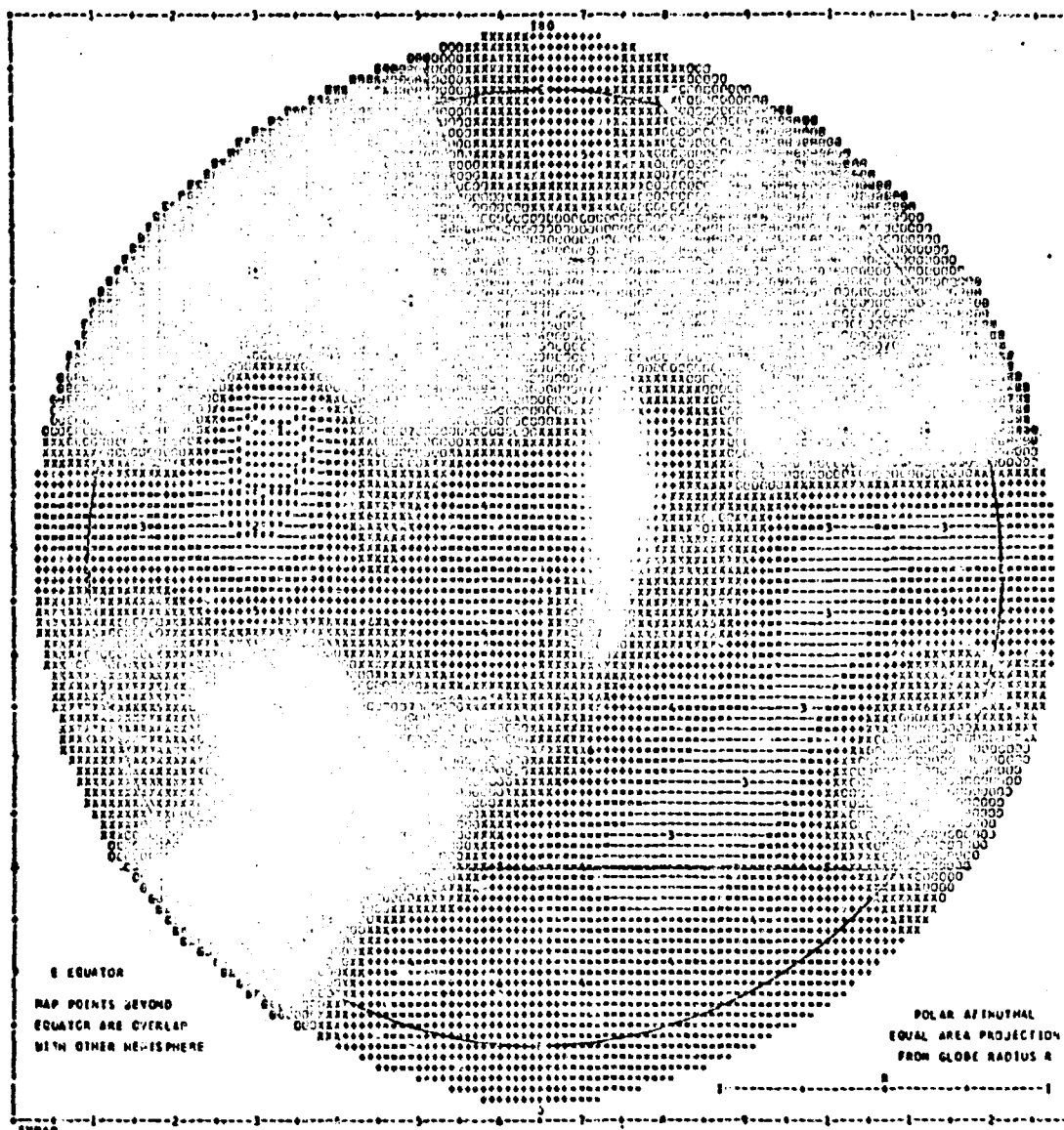function is drawn.

Figure 12d. The zero contour for the sunshine difference
function is drawn.

instructs the computer to do sophisticated interpolation at
every third character location in the horizontal direction
and every other character location in the vertical direction
and then uses simple linear interpolation to fill in the
gaps. Thus subtraction at every character location is
not necessary -- the result would be virtually indistin-
guishable from placing data points at every third location
and then interpolating between. In fact, the resulting
map would not differ significantly if even fewer data
points were used and this would save the trouble and
expense of writing and debugging a subroutine.

Thus, to refine the technique further, subtraction
at every character location was bypassed, although that
would certainly have worked and have been an improvement.
Instead, a new base map was layed out with almost three
times as many data points (a total of 201 including overlap
points) and an average spacing between adjacent data
points of about nine character locations. The equator
was the same size as before and was approximated by a
32 sided polygon. The series of overlap points beyond
the equator remained to insure continuity at the equator
and a series of data points lay on the equator itself
(on the previous base map there were no data points on
the equator). These improvements were designed to overcome
the omissions of previous difference mappings. Finally,
on this series of maps the 0° longitude line was rotated
by 180° to the top of the map and the symbolism ended
at the equator, not at the overlap points as on the
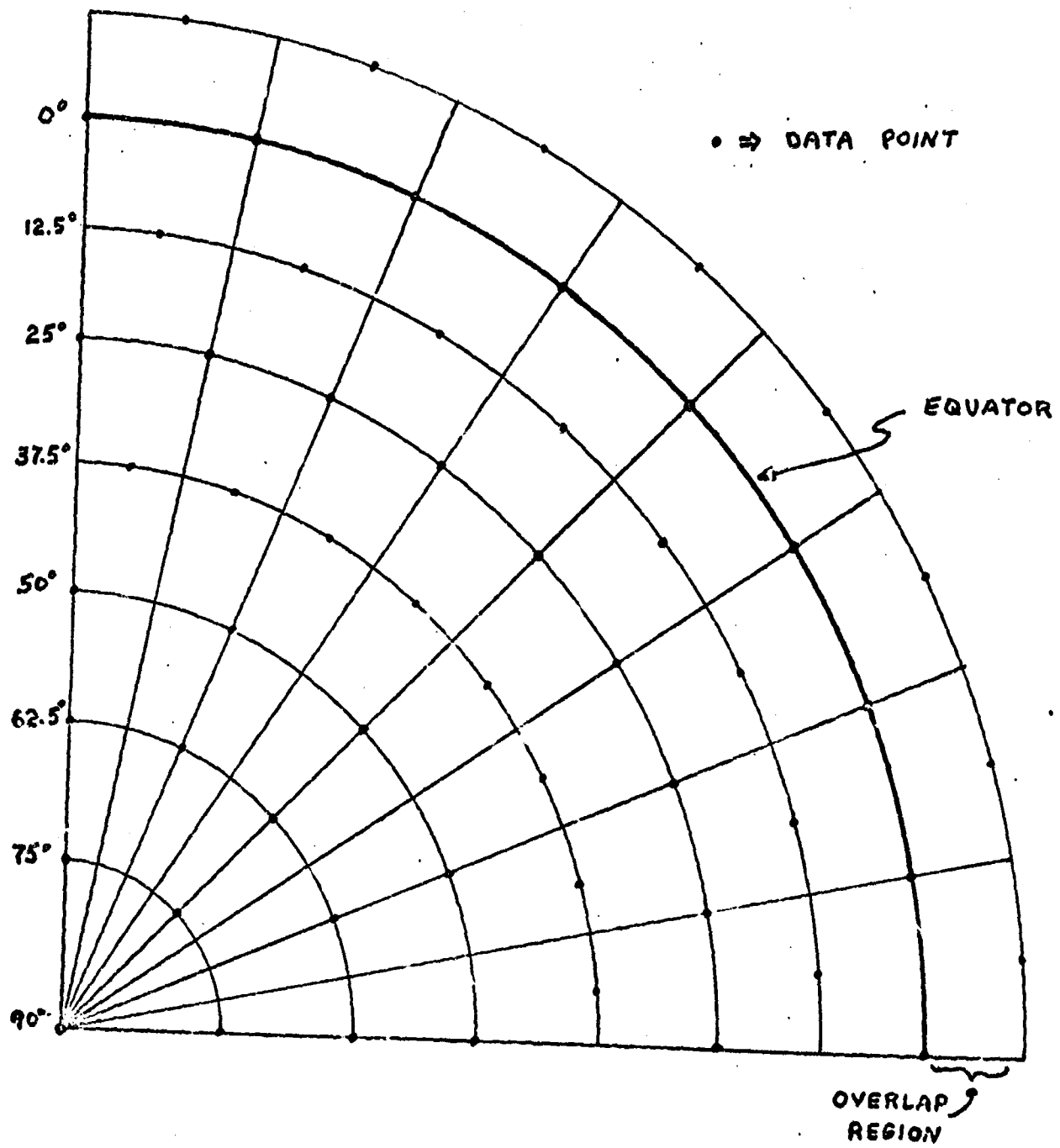previous maps.

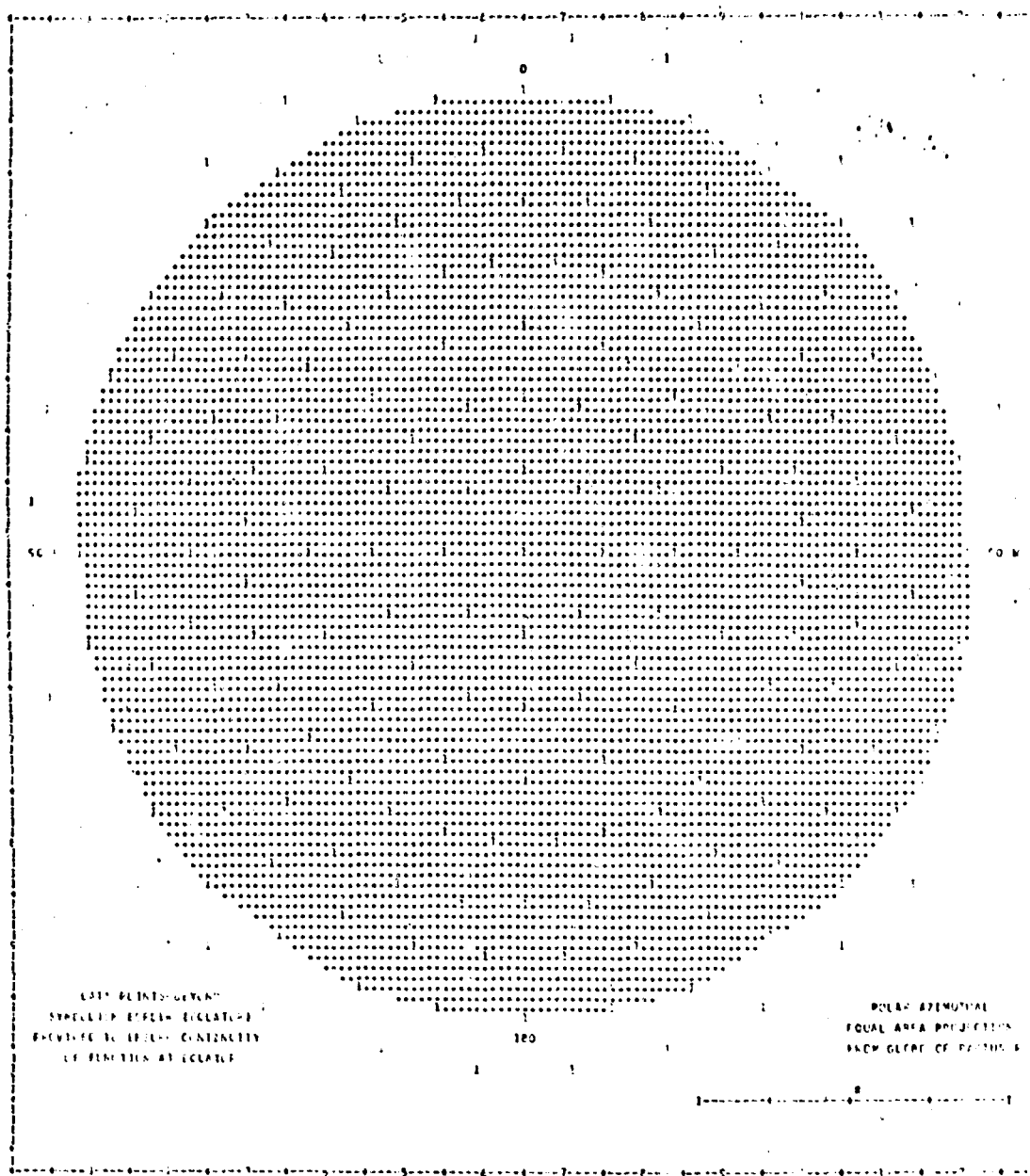Figure 13.   Data Point Locations on one Quadrant of Revised
Base Map

Figure 14. Revised Base Map for Mapping Difference
Functions (201 Data Points)

To eliminate the guesswork involved in plotting the 0 contour when it falls in the middle of a contour level, several of the electives available with SYMAP are utilized. These electives are used to make the computer actually print out the zero contour and thus eliminate another human task in the solution procedure. This is done by specifying a map of the difference function in two levels only, with the zero contour line, or locus of equal-valued sets of antipodal points, between them. Appropriate symbolism might be + signs for all regions where the difference functions is greater than zero and - signs for all regions less than zero. The darkest symbol is specified for the contour line to make it stand out.

These changes would produce a highly readable difference map and locate the 0 contour to within the accuracy of the original data. But, even more information can be obtained from the same map. Since there will be some error involved in estimating the function values on the original contour map, the spatial location of the 0 contour may not be located precisely. An error in estimating the original function value would show up as a small displacement of the 0 contour on the difference map. An estimate may be made of just how far the 0 contour might stray from its calculated position by putting limits on the values permitted in each contour level. If an a rage error in reading an Energy value is taken as $\pm$ 5 kcal/cm^2/year then the propagated error in the difference function is $\pm$ 7 kcal/cm^2/year. Thus,

if the range of the upper contour level is limited to +7
and the lower contour level to -7, the appropriate symbolism
will be printed only where the value of the difference
function is within those ranges. All other areas will be
blank. The computer output would consist of black 0
contour lines surrounded by bands of symbolism of various
widths within which the 0 contour might actually be located
if there were some error in estimating function values.

Using certain mapping electives to produce the
desired results, refined difference mappings were made of
the Energy and Sunshine functions (average propagated
error for the Sunshine difference mapping was estimated
at ± 140 hrs/yr). The results were excellent. In most
regions of the map the contour levels were wide enough to
permit the black zero contour to be printed. In other
regions where the contour levels were only one character
space wide, (A, Figure 15b), the   cise location of the
zero level was known (between adjacent + and - signs)
even though the black contour was omitted. Only where
the difference function was changing so quickly that
scattered symbolism appeared (B, Figure 15b) was there
some uncertainty about the exact position of the zero
contour, but even this uncertainty was small. Each time
the zero contour intersected the equator it reappeared at
the diametrically opposite point. This former problem
area was finally eliminated. Where previously the exact
behavior of the zero contour had been in doubt it was
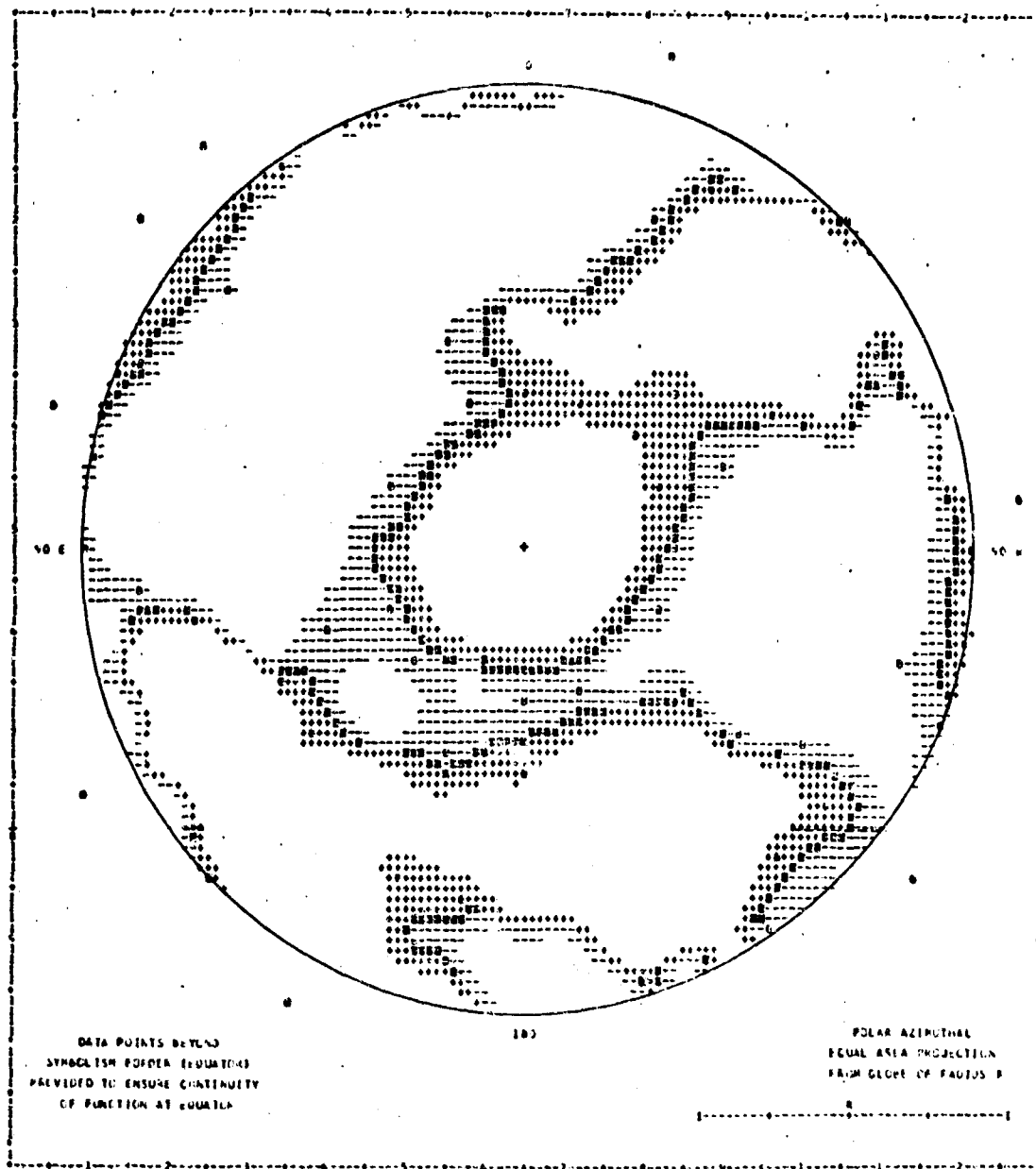precisely defined.

Figure 15a.   Revised Difference Function Mapping of
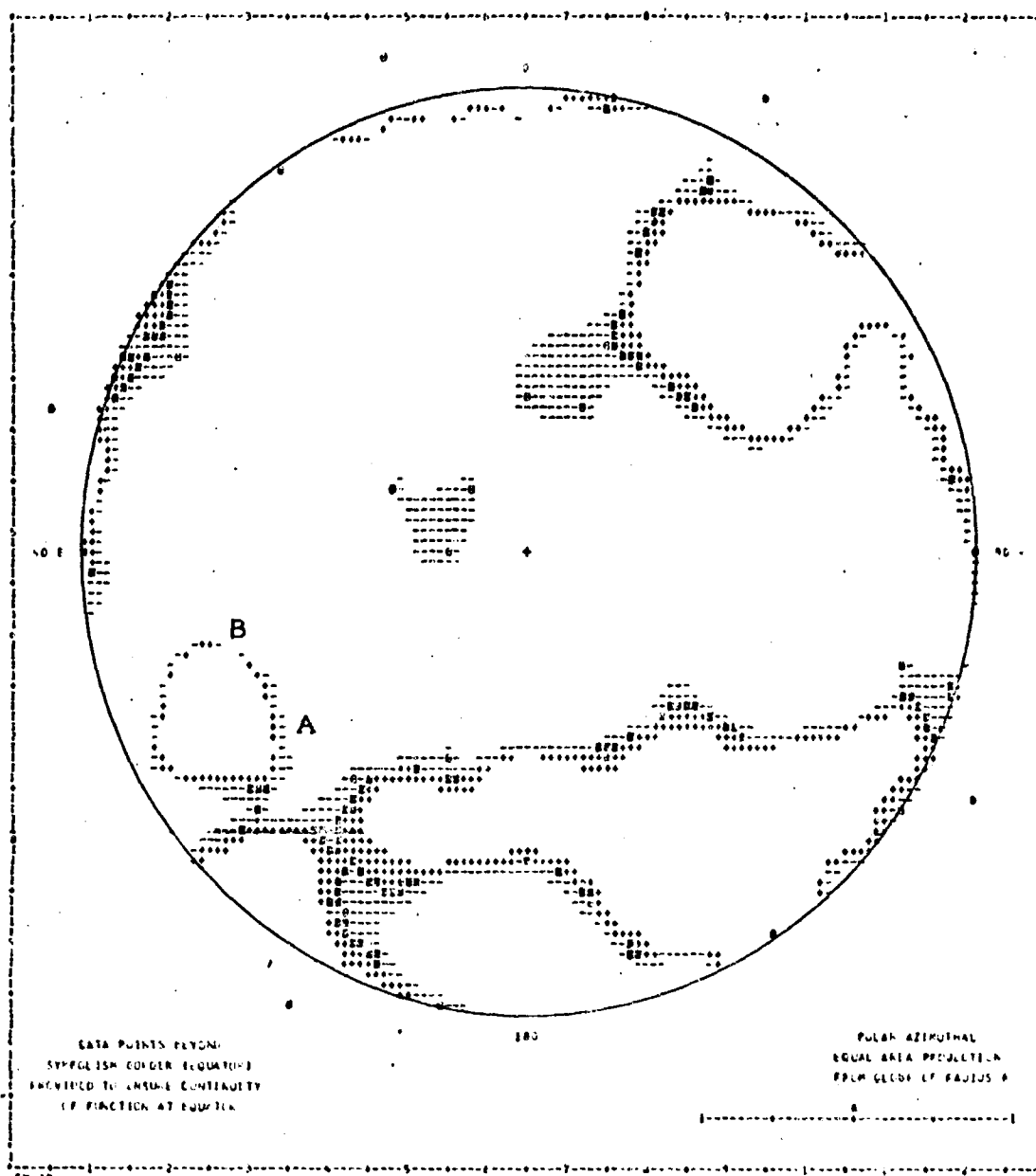
Energy Function

Figure 15b. Revised Difference Function Mapping of
Sunshine Function

The zero contour displays many of the properties of
the intersection of two surfaces. If the difference .
function is envisioned as a three-dimensional surface,
the zero contour may be obtained by intersecting the
surface with a level surface of zero value. Although
all the zero contours of both functions form a series
of closed curves (ignoring discontinuities at the equator)
the possibility of isolated points or line segments at
the zero level must not be overlooked. Under what
conditions, then, do two continuous surfaces intersect
in a point and in a line segment?

Point intersection may be interpreted as the limiting
stage of intersection between a level surface and a peak.
When the surface intersects the base of the peak, the
intersection is a closed curve. As the level surface is
moved upward, however, the closed curve shrinks in size
to a point and then vanishes (Figure 16a). It would appear
from this example that the condition for a single point
of intersection of the two function surfaces (i.e., the
locus of equal-valued sets of antipodal points is just
one set of points) is that the tangent level surfaces
to each function at that particular point be identical.
The case of the line segment can be evaluated in a similar
manner. In terms of intersecting surfaces, consider, for
example, a horizontal surface intersecting a ridge with
a horizontal ridgeline. As the level surface moves up
the ridge, the intersection shrinks from an initially
elongated closed curve to a line segment when the surface
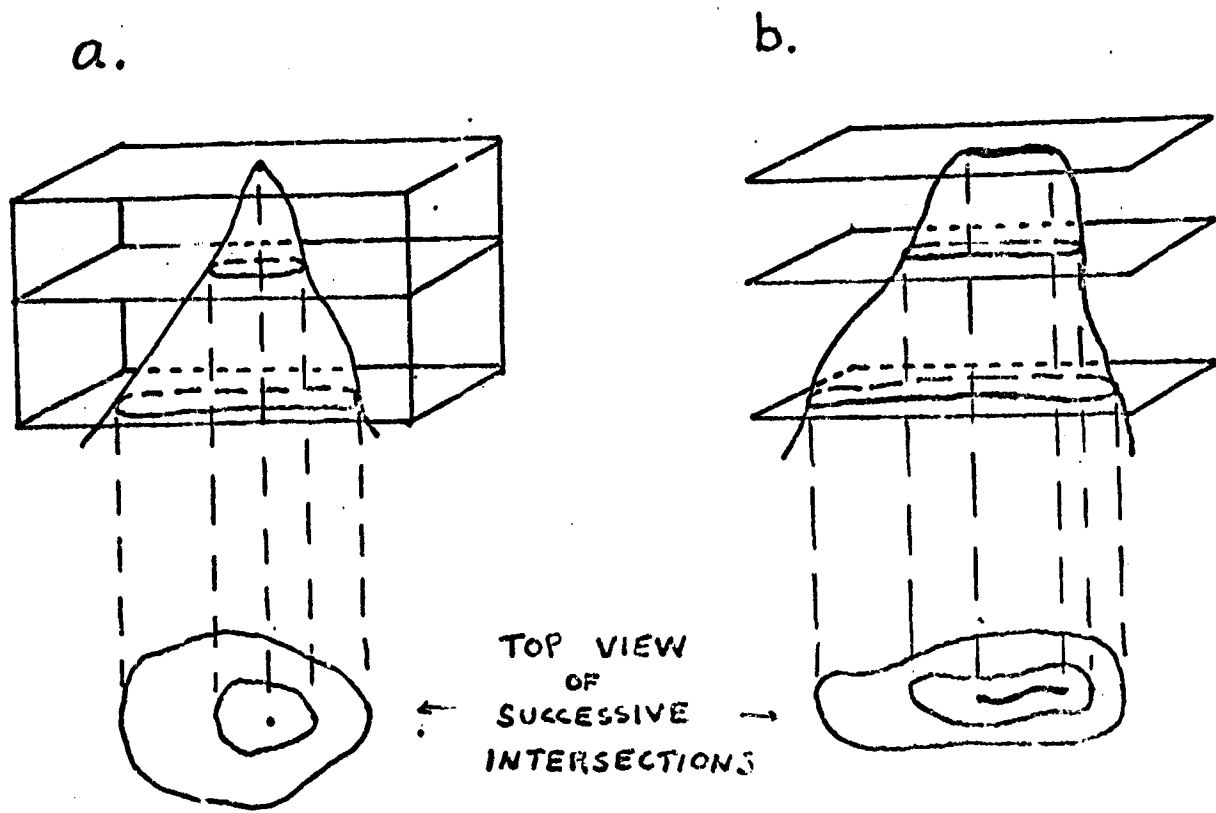is just tangent at the ridgeline. As previously mentioned,

a.

b.

TOP VIEW
OF
← SUCCESSIVE →
INTERSECTIONS

Figure 16. Intersecting Surfaces

neither the single point nor line segment loci of equal-
valued sets of antipodal points were observed on these
two difference function mappings.

The particular printout format of the difference
functions has yet another potential use. It provides
information which aids in answering the following question:
What happens to the spatial solutions of the existence
theorem as one or both of the functions is varied by
small amounts? As the first step in answering this
question, the spatial variation of the zero contour of
one of the difference functions under small perturbations
of the function is examined.

Only the case of small perturbations of the function
in a localized region is reviewed because larger variations
would result in changes too widespread and complex to
analyse. They would present, essentially, a completely
new function and could be treated as such, by mapping the
difference function, etc. The distinction between large
and small variations of the function is vague. It will
be convenient to consider small perturbations as those on
the order of the range of the contour levels of the
difference function. This implies, of course, that what
is considered small for one function may be large for
another, since the contour level ranges will vary consid-
erably among difference functions.

The computer printout of the zero contours for a
difference function yields much information on this prob-
lem. The total width (measured perpendicular to a

tangent to the contour at that point) of the two contour

levels at any particular point on the zero contour gives

a rough indication of the slope of the difference function

at that set of antipodal points. First, observe that the

change in the difference function (Δf) over the two contour

levels is just twice the propagated average error of the

original function (call this 2ϵ). An approximate

corresponding change in the domain of the function (Δx)

is determined by actual measurement of the width of the

two contour levels. (Note: the scale of the map is

constant in a circumferential direction but varies in

the radial direction.) Thus the gradient of the difference

function can be approximated at any point on the zero

contour by $\frac{\Delta f}{\Delta x}$.

To avoid complications which arise from the varying

scale of the difference map, the gradient of the difference

function will be discussed in qualitative terms only.

Where the gradient is small, (that is, where the width

of the two contour levels is large (A, Figure 17) a

small change in the value of the difference function at

some point on the zero contour would result in a relatively

large change in the spatial location of the zero contour.

For instance, if the function is varied so that the

difference function changes by "ϵ" at some point on the

zero contour, one would expect the zero contour to shift

to one side by about the width of the contour level.

The direction of the shift would depend upon whether the

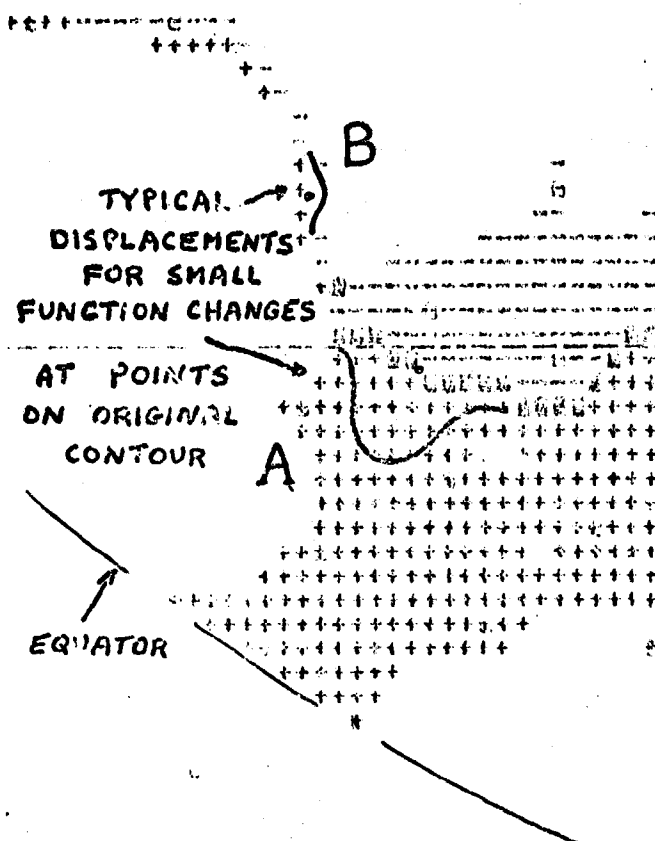difference function increased by ϵ or decreased by ϵ.

Figure 17. Displacement of the zero contour of the difference function for a small function change at some set of antipodal points varies with the inverse of the gradient at the points in question.

Similarly, in a region where the gradient is steep (B, Figure 17) the difference function is changing rapidly. If the difference function is changed by ϵ at some point on the zero contour in this region, the zero contour will experience a much smaller displacement (one character location or less).

Thus the displacement of the zero contour of the difference function as a result of small, localized function variations can be estimated by examining the printout as described above. These conclusions hold for changes in the difference function resulting from small spatial shifts of the original function as well as small variations in the original function values. The continuity of the function assures the equivalence of these two cases.

Only small changes in the difference function have been discussed because the primary concern here is with what happens to the zero contour in a small region about the point in question. A large change in the difference function at a given point would produce significant changes in the locality under consideration and the approximate techniques discussed previously would not give meaningful results. But, because the distinction between large and small changes is necessarily vague and arbitrary, it may at some time be necessary to approximate the displacement of the zero contour for difference function changes as large as 2ϵ or 3ϵ. This can be accomplished with only slight revisions in the

previous techniques. By changing the electives which control the number of contour levels and their ranges, a difference map could easily be produced with four levels, for instance, which would cover twice the range of previous maps. To use the Energy function as an example, the levels would range from 14-7, 7-0, 0-(-7), and (-7)-(-14). The displacement of the zero contour could then be estimated in the locality of a point where the Energy difference function changed by as much as 14 $kcal/cm^2/year$.

If the approximate displacements of the zero contours for small changes in both functions are known in some region, the behavior of an intersection of the zero contours, i.e., the pair of antipodal points at which each function has equal values, can then be inferred. The gradient of the difference function at any point on the zero contour line is perpendicular to the tangent to the contour at that point. Small changes in the difference function will displace the zero contour roughly parallel to its former position in the direction of either plus or minus the gradient at any particular point on the contour. Thus, to approximate the motion of the point of intersection of two zero contours, take the vector sum of the inverse of the gradient of each contour ($\frac{1}{\nabla f}$) at that point. (The direction is determined by the sign of the change.) The resultant vector indicates the instantaneous direction and magnitude of motion of the point of inter- section. An estimation of its displacement may be made by plotting the new locations of both zero contours and then noting the new position of their point of intersection, assuming they still intersect.

Points of intersection of zero contours may vanish
when the difference functions are varied slightly but
the theorem requires the existence of at least one point
of intersection regardless of how greatly the functions
change, as long as the changes are continuous. The
significance of the particular spatial locations of the
sets of antipodal points at which both of the functions
have equal values is a complex problem. There is an
unclear dependence on the degree to which the two functions
are related. The Energy and Sunshine functions, for
instance, are in some ways related (one would expect the
Energy function to be dependent on the Sunshine function)
and this is borne out by a superposition of the zero
contours of both difference functions. The zero contours
of these two difference functions intersect over twenty
times but this in itself is not significant because the
randomly generated functions previously developed had
over twenty intersections as well. What is more signif-
icant is the degree to which the zero contours of one
function seem to correspond to those of the other. In
many regions, the contours are nearly parallel but
slightly displaced -- in others they practically coincide.
These results were not totally unexpected.

As previously mentioned, Sunshine and Energy are
basically latitude dependent phenomena. Were it not for
the presence of atmospheric phenomena, which in turn are
in many ways dependent on other physical factors, the zero
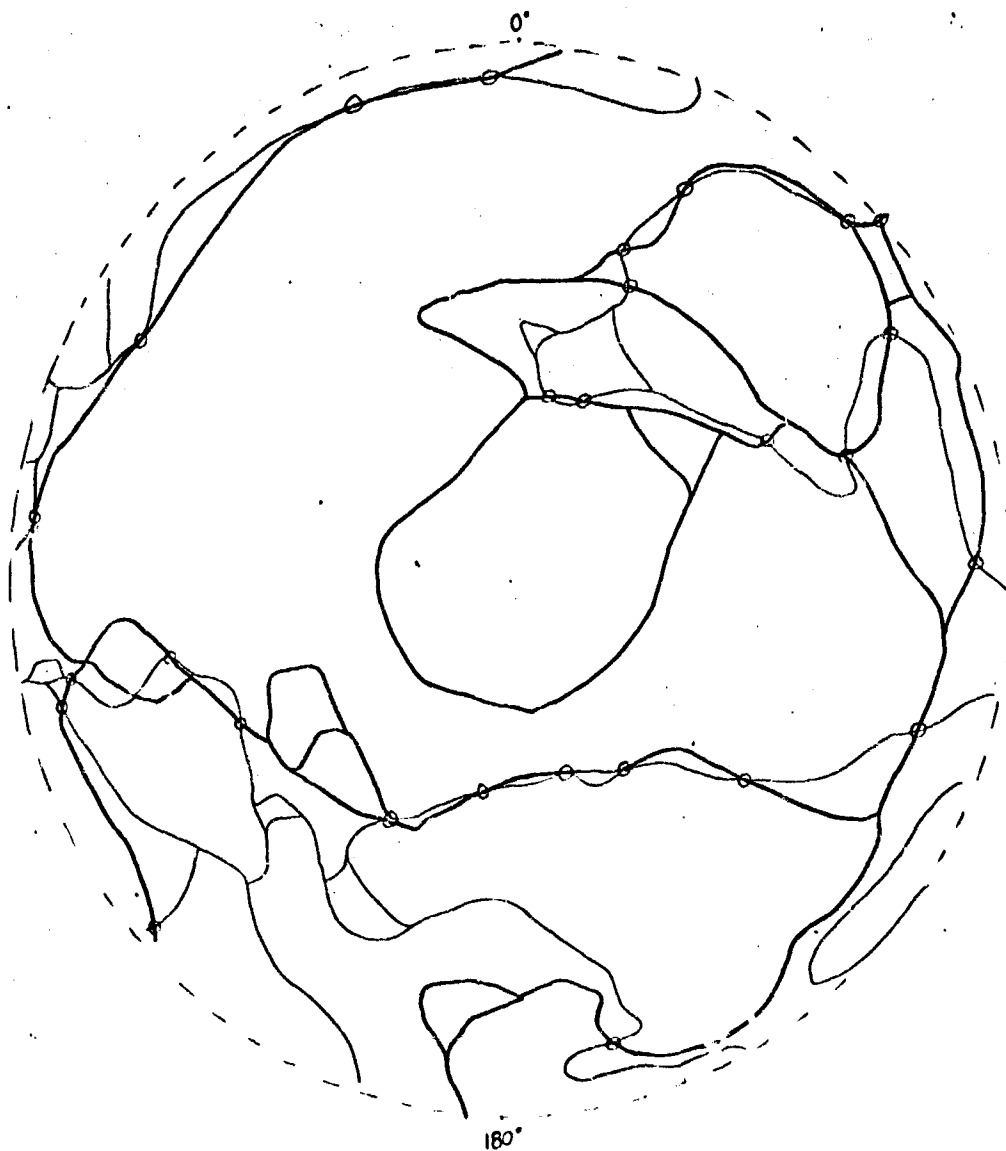contours of both difference functions would have covered

Figure 18a. Loci of Sets of Equal Valued Antipodal Points for Energy Function (Thick Lines) and Sunshine Function (Thin Lines)
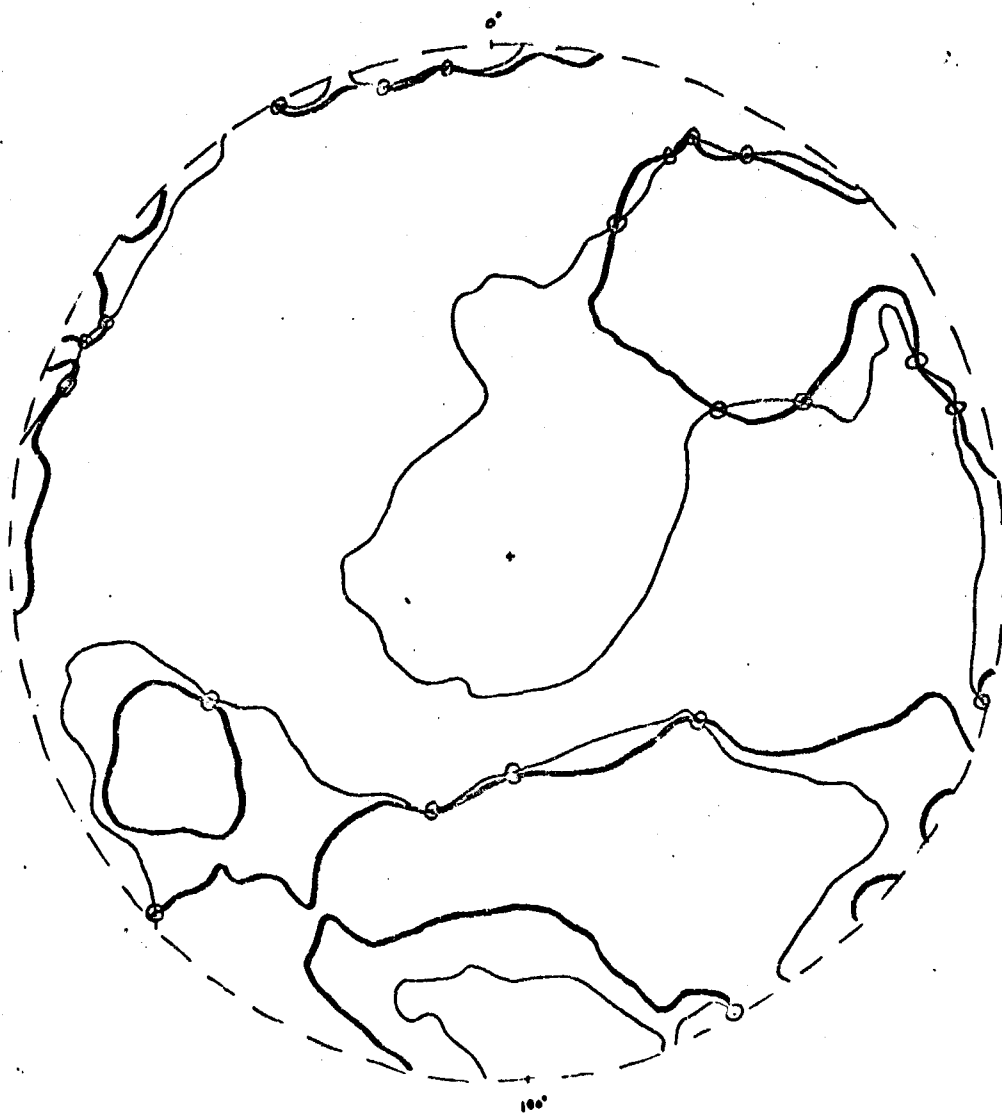
note: The intersections are circled.

Figure 18b. Zero Contours of Energy Difference Functions (Thin
Lines) and Sunshine Difference Functions (Thick Lines)

note: The intersections are circled.

the whole sphere. Where the zero contours do in fact exist
may indicate that the same atmospheric conditions and their
preconditions exist at those sets of antipodal points.

To follow up this hunch, a difference function was
mapped for a continentality potential, computed by Donald
Shepard. He divided the earth into trapezia of 5° x 5°
and assigned a value to each: 100 if the trapezoid con-
tained mainly land, 0 if it was mostly water. A potential
value was then calculated for each trapezoid by weighting
the value of every other trapezoid by the inverse of the
distance-squared. A contour line of value about 70 traces
out recognizable continents. Thus, the continentality
potential closely resembles a map of the earth, accen-
tuating lines of demarcation between water and land
masses.

Since the continentality potential is a continuous
function over the surface of a sphere the zero contours
of its difference function must meet the criteria pre-
viously established. Examination of the difference
function shows that it does. This difference function is
interesting because it has very little more than what was
determined as the minimum locus of sets of antipodal
points with equal value; that is, a curve connecting
two opposite points on the equator of the hemispherical
map. There are only two small additional closed curves.

Figures 20a and 20b show the zero contours of the
continentality potential difference function plotted on
the same hemispherical map as the Energy and Sunshine
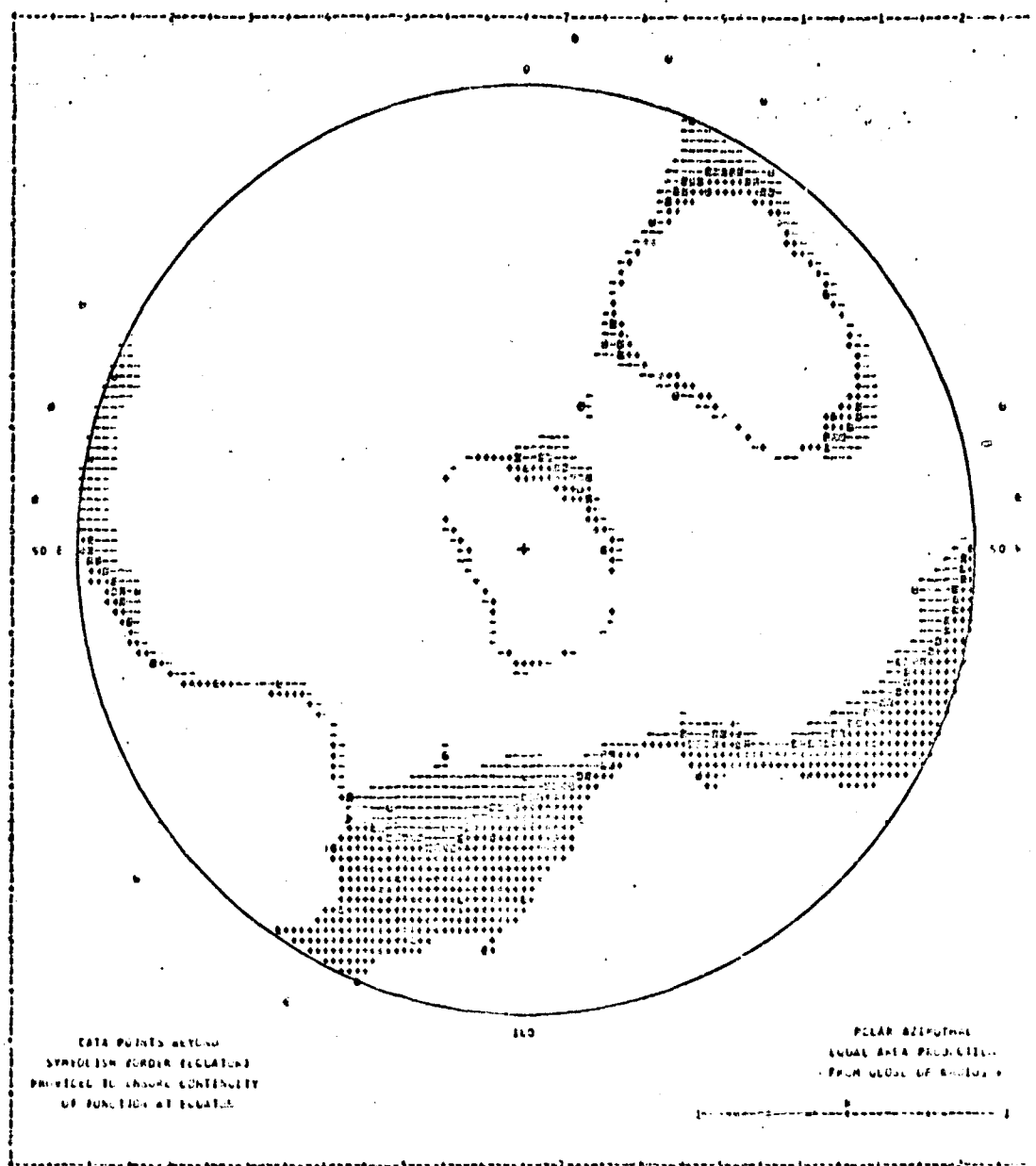
Figure 19. Revised Difference Function Mapping of Continentality Potential Function
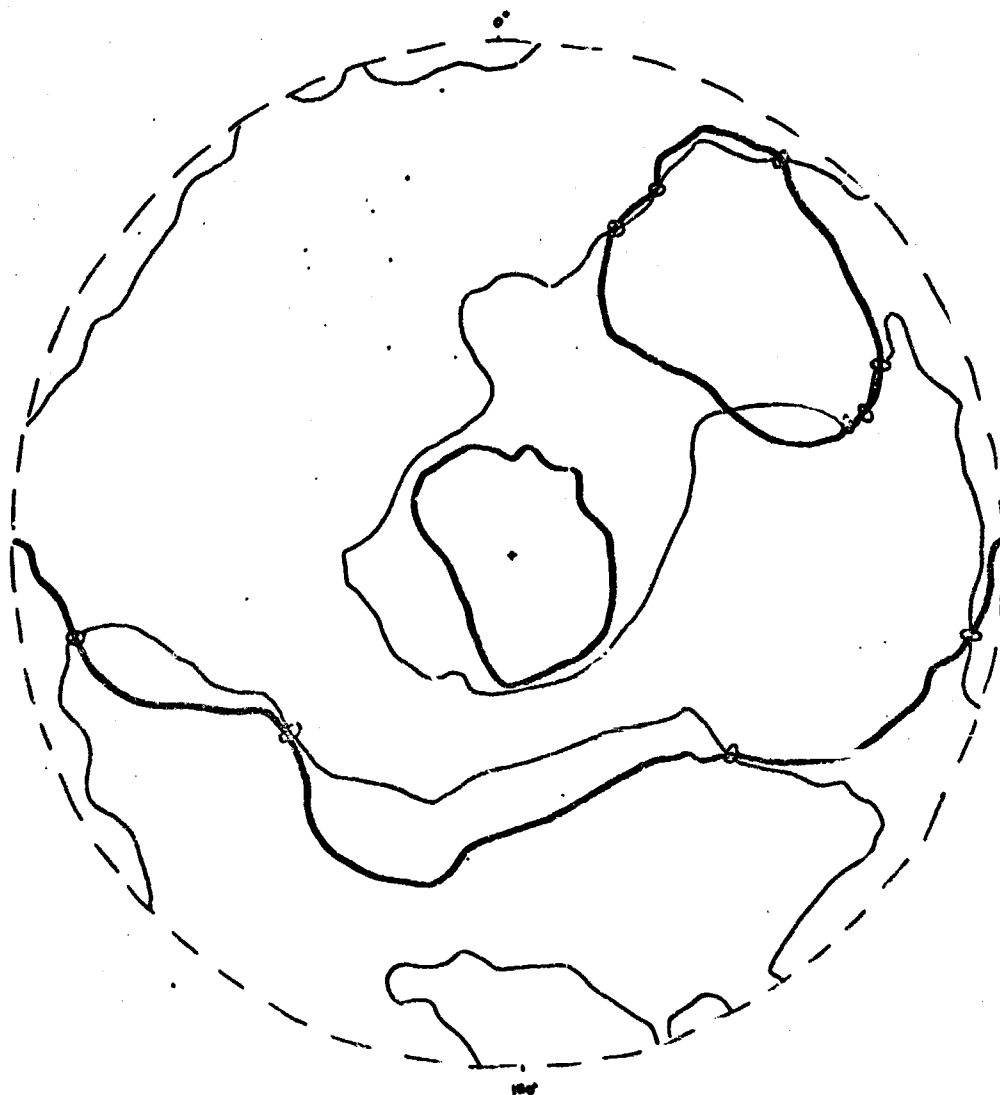
Figure 20a.   Zero Contours of Energy Difference Function
              (Thin Lines) and Continentality Potential
              Difference Function (Thick Lines)

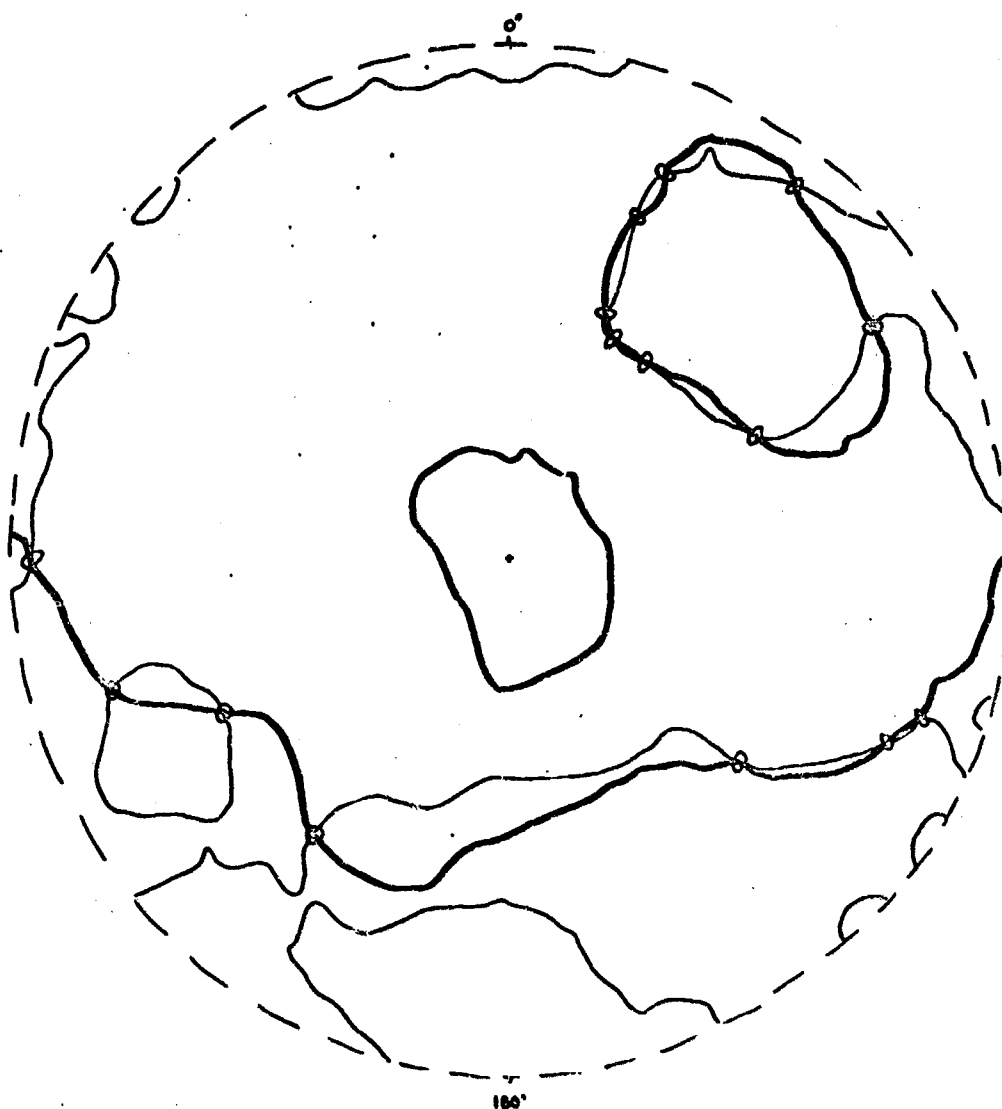   note: The intersections are circled.

Figure 20b. Zero Contours of Sunshine Difference Function (Thin
Lines) and Continentality Potential Difference Function
(Thick Lines)

note: The intersections are circled.

difference functions. Although there are few points
of intersection, large segments of the continentality
zero contours closely approximate the zero contours
of both the Energy and Sunshine difference functions.
This would seem to imply that there is some correlation
between the factors that make the Energy and Sunshine
functions have equal values at antipodal points and
the degree of continentality and marine influence at
these sets of antipodal points.

Furthermore, if the geographical locations of
the zero contours of these three difference functions
are examined, there appears to be more than coinciden-
tal agreement with the land-water boundaries on the
globe. This suggests that those atmospheric phenomena
which make the Energy and Sunshine functions equal at
sets of antipodal points are the result of physical
conditions related to the influence of land and water
masses at the boundaries between them. Couched in such
general terms, this proposition may be a statement of
the obvious, but the insights developed by this par-
ticular method of analysis might produce valuable
results with more intensive study.

It is difficult, if not impossible, to assign great
significance to the geographical location of the inter-
section of zero contours of two functions if the functions
are totally unrelated. The locus of intersections would
be little more than a mathematical curiosity. An attempt
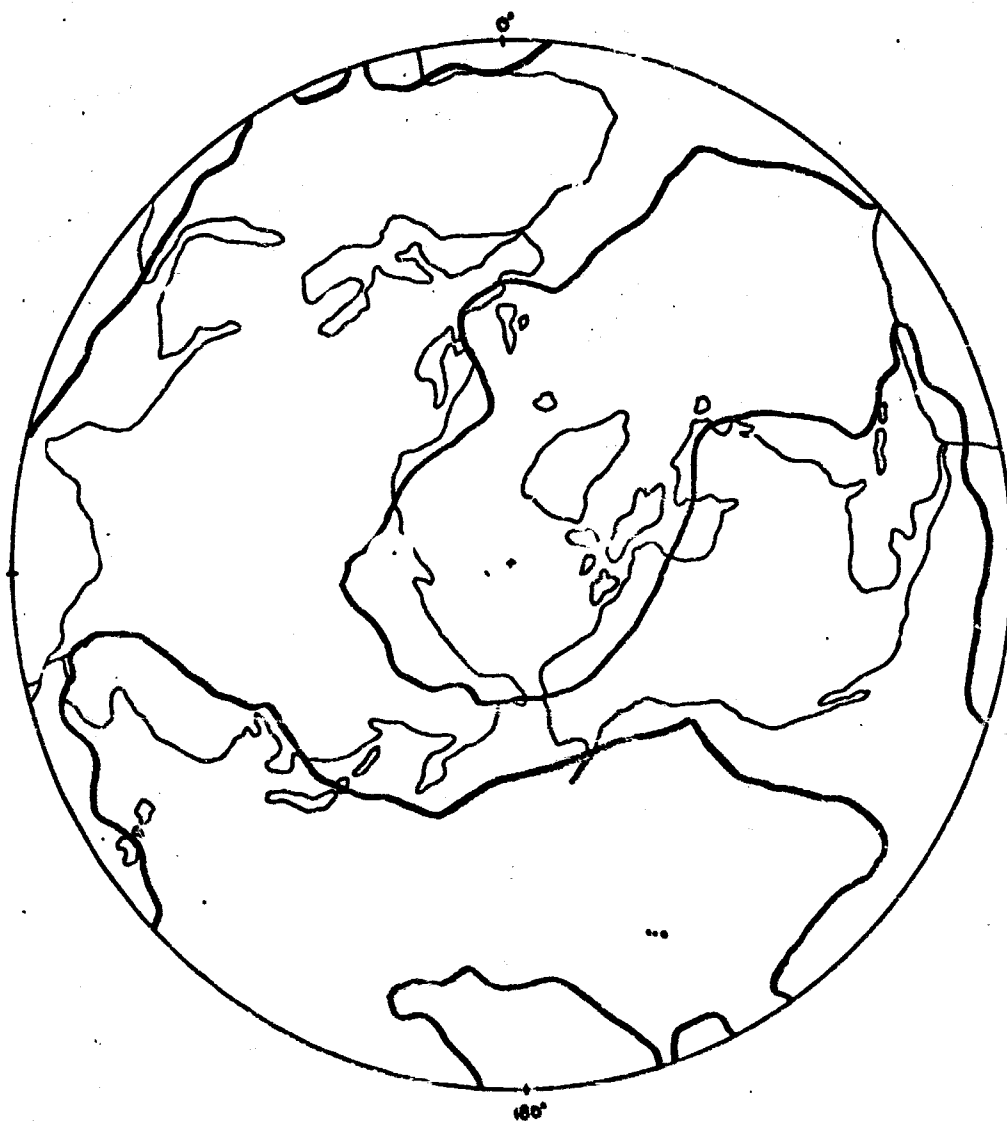to fabricate causal relations among physical phenomena

Figure 21a. The zero contours of energy difference function
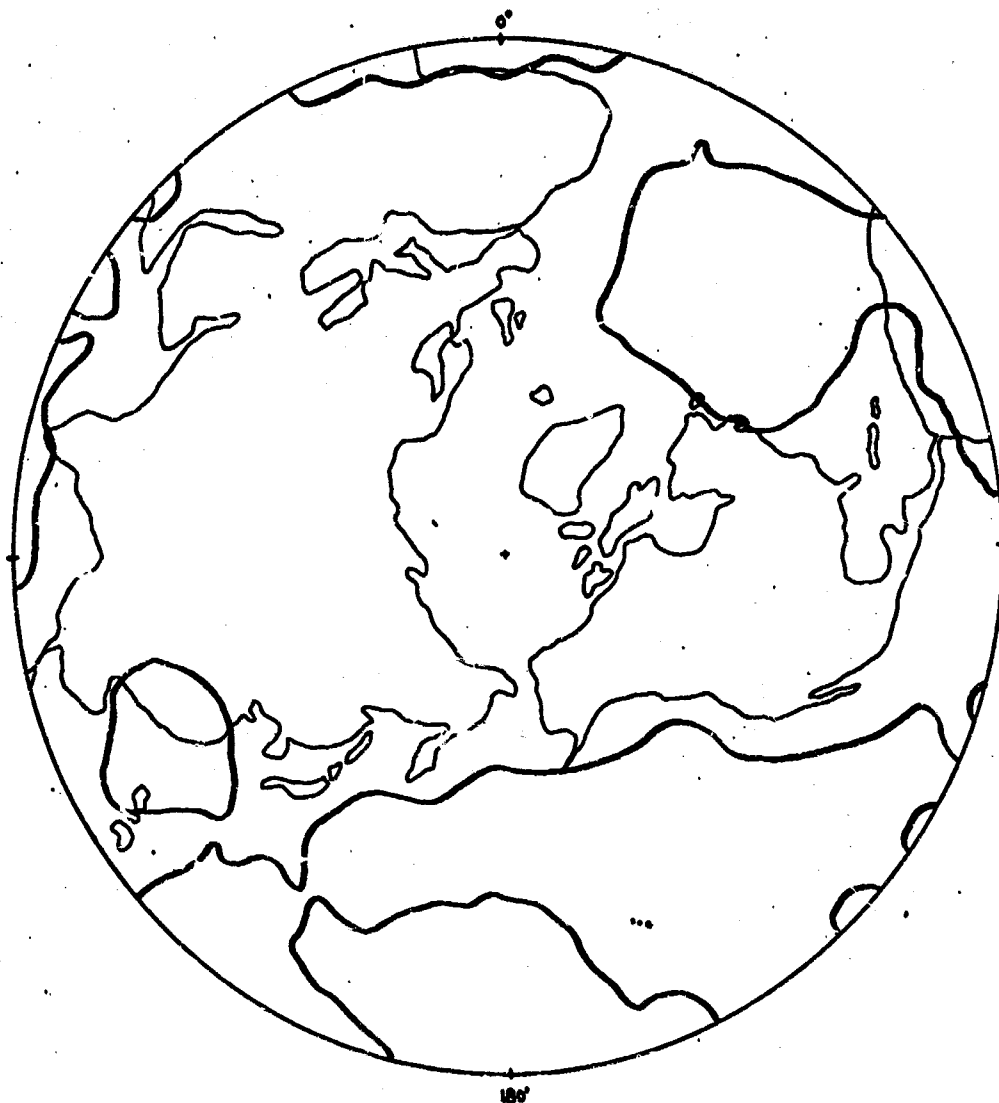are here plotted on the Northern Hemisphere.

Figure 21b.   The zero contours of sunshine difference function
are plotted on the northern hemisphere.

Figure 21c. The zero contours of continentality potential
difference function are plotted on the northern
hemisphere.

to explain the intersections of totally unrelated functions might be an exercise in futility although, of course, the results might be suggestive of patterns of relationships, hitherto ignored. Caution should be exercised when trying to explain the intersection of zero contours of functions which are related in an unclear way. The mathematical necessity of at least one point of intersection limits the significance one should attach to any explanation which ascribes to other factors, such as geological, climatic, etc., the responsibility for the points of intersection. These other factors do play a role, perhaps, in determining the number of intersections and their locations, but this depends to a large extent upon the nature of the particular functions under consideration. On the other hand theories ignoring the necessity of these intersections are similarly suspect.

The uncertainties and inaccuracies resulting from repeated manipulation of the original data have been continually stressed. In judging the reliability of the final results, one must also consider the inherent limitations in the use of the mapping program. It is critically important that the spacing of the data points be finer than the scale of the phenomena which is being mapped. This problem plagued the first attempt to construct a difference function. The data point spacing was so wide that two adjacent data points might be in positive regions, for instance, while a negative region between them was missed, resulting in the omission of a zero

contour line. This situation was rectified by the addition
of more datapoints. There could conceivably be a function
for which the addition of more data points would still
not solve this problem. A detailed function showing
population potential of the world would have "spikes" at
most city locations. But even a reasonably sized SYMAP
contour map with data points at every character location
would probably miss a great many of the city spikes. If
the intention were to show the general character of the
potential then the map would have sufficed, but if it
were important to show the precise function value at
city locations, a considerably larger map would have had
to be used.

The solution procedure advocated in the preceding
pages for finding spatial solutions for the original
existence theorem is not necessarily the best solution or
a final solution in any sense. It is, however, a relatively
quick, accurate and inexpensive computer solution to the
problem.

To review briefly, a base map was constructed from
polar azimuthal equal area projections of hemispheres of
a globe such that each point on the map represented a
pair of points on a globe -- a given point in one
hemisphere and its antipodal point in the other hemis-
phere. (Antipodal points were aligned by a simple rota-
tion of one of the hemispheres.) The base map contained
a large number of closely spaced data points and a set
of points beyond the equator to insure the continuity of

the function across the equator. The "difference function"
of the function under consideration is plotted on this
base map. The value of the difference function at a
given set of antipodal points is computed by subtracting
the value of the function at the appropriate point in
the northern hemisphere from the value at its antipodal
point in the southern hemisphere. The SYMAP program is
then used to produce a contour map of the difference
function. The zero values of the difference function
are of prime importance since at these sets of antipodal
points, the original function will have equal values.
Using the appropriate electives in the SYMAP program, the
computer is instructed to print out a black contour line
at the value zero and one contour level on each side of
the zero contour. The range of the contour levels is
determined by estimating the average error in assigning
function values at data points. Where the difference
function changes so rapidly that there are insufficient
character locations to print the black zero contour,
it must lie between adjacent positive and negative valued
points. The two contour levels not only provide an
estimate of the possible inaccuracies in the location of
the zero contour but also give an indication of the
gradient of the difference function. This allows evalua-
tion of the displacement of the zero contour for varia-
tions in the original function whi   esult in small
changes in the difference function.

To find where both functions each have equal values
at sets of antipodal points, the zero contours of the

difference functions of the two original functions were
drawn on the same map and the points of intersection were
marked. These are the desired points. The displacements
of the points of intersection resulting from variations
in both original functions can be estimated by noting the
displacement in both zero contours and finding the new
point of intersection (assuming they still intersect
after having been displaced).

Comparison of results obtained by this method and
by the tedious process of hand-tracing and overlapping
the original contours showed that none of the zero contours
had been omitted but that many ambiguities and uncertain-
ties inherent in the tracing and overlapping method had
been eliminated. Improvements in this procedure can
still be made. If the accuracy of the data be sufficiently
good, function values could be subtracted at every char-
acter location rather than only at data points. As
stated previously, however, because of the nature of
the interpolation process, this method would not produce
significant improvements. The use of more sophisticated
computing equipment which produces line drawings of three-
dimensional surfaces on paper or cathode ray tubes might
vastly improve comprehension. Superposition of function
surfaces or intersection of the difference function surface
with a zero level surface would be possible using equip-
ment of this nature and would produce results similar to
what has been obtained but in a more graphic and versatile
manner.

Even this relatively simple computer assistance in finding spatial solutions to the given existence theorem demonstrates the possibility of using computer mapping programs to find spatial solutions to that whole class of existence theorems having explicit or implicit spatial connotations. After initial problems related to adopting the particular existence theorem to a computer solution are resolved, computer mapping techniques are capable of quick and accurate spatial solutions with a minimum of human manipulation. This is another indication that the problem solving capabilities of computer mapping are at least as diverse and effective as are its proven capacities graphically to represent stores of spatially ordered information. These capabilities lie largely untapped, waiting to be exploited.

DOCUMENT CONTROL DATA · R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
President and Fellows of Harvard College	2b. GROUP

3. REPORT TITLE GEOGRAPHY AND AN EXISTENCE THEOREM -- A Cartographic Computer Solution to the Localization on a Sphere of Sets of Equal-Valued Antipodal Points for Two Continuous Distributions with Practical Applications to the Real Earth

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Interim Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Stephen E. Selkowitz

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
5 August 1968	64	0

9a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
ONR 00014-67A-0298-0004	
b. PROJECT NO.	ONR Technical Report #21
NR 389-147	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	Harvard Theoretical Geography Paper #21

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research, Geography Branch

13. ABSTRACT

Computer mapping is used here to effect the "solution" to the problem of locating on the earth's surface those sets of antipodal points for which equal values occur for some function f and equal values occur for some function g when both of these functions may be regarded as continuous over a sphere. An existence theorem in mathematics (proposed by Ulam and proved by Borsuk) states that at least one such set of antipodal points must always exist. Geographers, naturally enough, are interested in finding the locations of such sets for pertinent real phenomena on the earth's surface.